

#2

Attorney Docket No. 1341.1119

JC903 US62 Pro
10/050862
01/18/02

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent Application of:

Takanori UGAI, et al.

Application No.:

Group Art Unit:

Filed: January 18, 2002

Examiner:

For: INFORMATION USE FREQUENCY PREDICTION PROGRAM, INFORMATION USE FREQUENCY PREDICTION METHOD, AND INFORMATION USE FREQUENCY PREDICTION APPARATUS

**SUBMISSION OF CERTIFIED COPY OF PRIOR FOREIGN
APPLICATION IN ACCORDANCE
WITH THE REQUIREMENTS OF 37 C.F.R. § 1.55**

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

In accordance with the provisions of 37 C.F.R. § 1.55, the applicant(s) submit(s) herewith a certified copy of the following foreign application:

Japanese Patent Application No. 2001-153116

Filed: May 22, 2001

It is respectfully requested that the applicant(s) be given the benefit of the foreign filing date(s) as evidenced by the certified papers attached hereto, in accordance with the requirements of 35 U.S.C. § 119.

Respectfully submitted,

STAAS & HALSEY LLP

By:

James D. Halsey, Jr.
Registration No. 22,729

Date: January 18, 2002

700 11th Street, N.W., Ste. 500
Washington, D.C. 20001
(202) 434-1500

JC903 U.S. PRO
10/050862
01/18/04



日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日

Date of Application:

2001年 5月 22日

出願番号

Application Number:

特願2001-153116

出願人

Applicant(s):

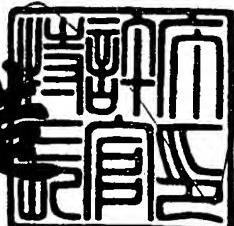
富士通株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年10月 3日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



【書類名】 特許願

【整理番号】 0151174

【提出日】 平成13年 5月22日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 15/40

【発明の名称】 情報利用頻度予測プログラム、情報利用頻度予測装置および情報利用頻度予測方法

【請求項の数】 10

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 鵜飼 孝典

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 三末 和男

【特許出願人】

【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【代理人】

【識別番号】 100089118

【弁理士】

【氏名又は名称】 酒井 宏明

【手数料の表示】

【予納台帳番号】 036711

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

特2001-153116

【物件名】 要約書 1

【包括委任状番号】 9717671

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 情報利用頻度予測プログラム、情報利用頻度予測装置および情報利用頻度予測方法

【特許請求の範囲】

【請求項1】 コンピュータを、

第1の時系列情報の利用頻度に関する時間的変化を表す第1のパターンと、第2の時系列情報の利用頻度に関する時間的変化を表す第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行う時間的操作手段、

前記所定時間単位毎に、前記第1の時系列情報と前記第2の時系列情報との相関係数を算出する相関係数算出手段、

前記相関係数算出手段により算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組を特定する組特定手段、

前記組に対応する第2のパターンに基づいて、前記組を構成する第1の時系列情報の利用頻度を予測する予測手段、

として機能させることを特徴とする情報利用頻度予測プログラム。

【請求項2】 前記時間的操作手段は、第1のグループに属する複数の第1の時系列情報と、第2のグループに属する複数の第2の時系列情報との全ての組み合わせに関して、前記時間的な操作を行うことを特徴とする請求項1に記載の情報利用頻度予測プログラム。

【請求項3】 コンピュータを、前記予測手段における複数の予測結果に対して、前記利用頻度をキーとしてソートをかけるソート手段として機能させることを特徴とする請求項2に記載の情報利用頻度予測プログラム。

【請求項4】 前記時間的操作手段は、前記第1の時系列情報を基準として、前記第2の時系列情報を時間軸上で、順次、所定時間単位でシフトさせることを特徴とする請求項1～3のいずれか一つに記載の情報利用頻度予測プログラム

【請求項5】 前記時間的操作手段は、前記第1の時系列情報を基準として、前記第2の時系列情報を時間的に、所定伸縮単位で、順次、伸縮させることを

特徴とする請求項1～3のいずれか一つに記載の情報利用頻度予測プログラム。

【請求項6】 前記時間的操作手段は、前記第1の時系列情報を基準として、前記第2の時系列情報を時間軸上で、順次、所定時間単位でシフトさせ、シフトされた前記第2の時系列情報を時間的に、所定伸縮単位で、順次、伸縮させることを特徴とする請求項1～3のいずれか一つに記載の情報利用頻度予測プログラム。

【請求項7】 前記第1の時系列情報および前記第2の時系列情報は、インターネット上のキーワード検索エンジンにおけるキーワードの利用頻度の時系列情報であることを特徴とする請求項1～6のいずれか一つに記載の情報利用頻度予測プログラム。

【請求項8】 前記第1の時系列情報と前記第2の時系列情報とは、異なる収集経路を介して収集されることを特徴とする請求項1～7のいずれか一つに記載の情報利用頻度予測プログラム。

【請求項9】 第1の時系列情報の利用頻度に関する時間的変化を表す第1のパターンと、第2の時系列情報の利用頻度に関する時間的変化を表す第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行う時間的操作手段と、

前記所定時間単位毎に、前記第1の時系列情報と前記第2の時系列情報との相関係数を算出する相関係数算出手段と、

前記相関係数算出手段により算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組を特定する組特定手段と、

前記組に対応する第2のパターンに基づいて、前記組を構成する第1の時系列情報の利用頻度を予測する予測手段と、

を備えたことを特徴とする情報利用頻度予測装置。

【請求項10】 第1の時系列情報の利用頻度に関する時間的変化を表す第1のパターンと、第2の時系列情報の利用頻度に関する時間的変化を表す第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行う時間的操作工程と、

前記所定時間単位毎に、前記第1の時系列情報と前記第2の時系列情報との相関係数を算出する相関係数算出工程と、

前記相関係数算出工程で算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組を特定する組特定工程と、

前記組に対応する第2のパターンに基づいて、前記組を構成する第1の時系列情報の利用頻度を予測する予測工程と、

を含むことを特徴とする情報利用頻度予測方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、例えば、インターネットにおける各種情報（検索キーワード等）の利用頻度の予測に用いられる情報利用頻度予測プログラム、情報利用頻度予測装置および情報利用頻度予測方法に関するものであり、予測精度を高めることができる情報利用頻度予測プログラム、情報利用頻度予測装置および情報利用頻度予測方法に関するものである。

【0002】

従来より、インターネットでは、キーワード入力により検索エンジンで、当該キーワードが含まれる情報を検索することが可能である。ここで入力されるキーワードは、流行や世相を反映していることが多い。従って、各企業では、検索ログデータとしてのキーワードを有効に活用する方法を研究している。例えば、企業では、時系列分析の手法を検索ログに適用して将来的なキーワードの利用頻度を予測した後、この予測結果を営業活動に利用している。

【0003】

【従来の技術】

近時、インターネット上に多くの電子商取引サイトが開設されており、商品の販売が積極的に行われている。この電子商取引サイトでは、仕入れ管理や在庫管理が売り上げに直結する重要なファクタとなっている。例えば、一週間後の売れ筋商品を予測できた場合には、的確な仕入れ、無駄が少ない在庫管理を行うこと

ができ、売り上げに直結する。

【0004】

また、多くの電子商取引サイトでは、キーワード入力による商品等の検索サービスを提供しており、検索ログデータ（キーワード）を時系列分析し、売れ筋商品の予測を行っていると考えられる。例えば、「カメラ」というキーワードの場合には、検索ログデータにおける過去の「カメラ」の利用頻度を時系列分析し、過去の傾向から未来の傾向を予測する。

【0005】

【発明が解決しようとする課題】

ところで、前述したように、時系列分析によりキーワードの利用頻度を予測する手法では、高い精度を期待できない。すなわち、インターネット上のキーワードは、流行、世相等の世の中の流れに非常に敏感である。従って、同一のキーワードに関して、過去の傾向がそのまま未来の傾向に当てはまるとは限らないため、精度が低いと考えられる。

【0006】

本発明は、上記に鑑みてなされたもので、情報の利用頻度を高精度で予測することができる情報利用頻度予測プログラム、情報利用頻度予測装置および情報利用頻度予測方法を提供することを目的とする。

【0007】

【課題を解決するための手段】

上記目的を達成するために、本発明は、コンピュータを、第1の時系列情報の利用頻度に関する時間的変化を表す第1のパターンと、第2の時系列情報の利用頻度に関する時間的変化を表す第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行う時間的操作手段、前記所定時間単位毎に、前記第1の時系列情報と前記第2の時系列情報との相関係数を算出する相関係数算出手段、前記相関係数算出手段により算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組を特定する組特定手段、前記組に対応する第2のパターンに基づいて、前記組を構成する第1の時系列情報の利用頻度を予測する予測手段として機能させることを特徴とす

る。

【0008】

また、本発明は、第1の時系列情報の利用頻度に関する時間的変化を表す第1のパターンと、第2の時系列情報の利用頻度に関する時間的変化を表す第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行う時間的操作手段と、前記所定時間単位毎に、前記第1の時系列情報と前記第2の時系列情報との相関係数を算出する相関係数算出手段と、前記相関係数算出手段により算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組を特定する組特定手段と、前記組に対応する第2のパターンに基づいて、前記組を構成する第1の時系列情報の利用頻度を予測する予測手段とを備えたことを特徴とする。

【0009】

また、本発明は、第1の時系列情報の利用頻度に関する時間的変化を表す第1のパターンと、第2の時系列情報の利用頻度に関する時間的変化を表す第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行う時間的操作工程と、前記所定時間単位毎に、前記第1の時系列情報と前記第2の時系列情報との相関係数を算出する相関係数算出工程と、前記相関係数算出工程で算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組を特定する組特定工程と、前記組に対応する第2のパターンに基づいて、前記組を構成する第1の時系列情報の利用頻度を予測する予測工程とを含むことを特徴とする。

【0010】

かかる発明によれば、第1の時系列情報に対応する第1のパターンと、第2の時系列情報に対応する第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行い、所定時間単位毎に、第1の時系列情報と第2の時系列情報との相関係数を算出し、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組に対応する第2のパターンに基づいて、組を構成する第1の時系列情報の利用頻度を予測しているため、従来のように一つの時系列情報に基づいて、予測を行う場合に比して、情報の利用頻度を高精度で予測するこ

とができる。

【0011】

【発明の実施の形態】

以下、図面を参照して本発明にかかる情報利用頻度予測プログラム、情報利用頻度予測装置および情報利用頻度予測方法の実施の形態1～3について詳細に説明する。

【0012】

ここで、実施の形態1～3は、例えば、地震が発生し、インターネット上で「地震」というキーワードが多く利用されると、しばらくして、地震との相関関係が深い「保険」という別のキーワードが、「地震」の利用頻度と同じようなパターンで利用されるという点に着目したものである。すなわち、実施の形態1～3では、あるキーワードの利用頻度を、該キーワードと相関関係が深い別のキーワードの時系列データを用いて予測する点に特徴がある。

【0013】

（実施の形態1）

図1は、本発明にかかる実施の形態1の構成を示すブロック図である。この図に示した情報利用頻度予測装置100において、読込部101は、第1の検索ログデータA10および第2の検索ログデータA20を読み込む機能を備えている。第1の検索ログデータA10は、例えば、キーワード入力によりインターネット上の情報を検索する第1の検索エンジン（図示略）の検索ログデータである。

【0014】

具体的には、第1の検索ログデータA10は、図2に示したように、検索用のキーワードが入力された「日付」と、「時刻」と、入力された「キーワード」という時系列のデータから構成されている。同図に示した例では、2000/11/01 01:10:00に「カメラ」というキーワードが第1の検索エンジンに入力され、この第1の検索エンジンで「カメラ」をキーとしてインターネット上の情報が検索されたことを意味している。

【0015】

図1に戻り、第2の検索ログデータA20は、例えば、キーワード入力により

インターネット上の情報を検索する第2の検索エンジン（図示略）の検索ログデータである。具体的には、第2の検索ログデータA20は、図3に示したように、検索用のキーワードが入力された「日付」と、「時刻」と、入力された「キーワード」という時系列のデータから構成されている。

【0016】

同図に示した例では、2000/11/01 11:10:50に「年賀状」というキーワードが第2の検索エンジンに入力され、この第2の検索エンジンで「年賀状」をキーとしてインターネット上の情報が検索されたことを意味している。

【0017】

図1に戻り、第1の集計部102は、読み込部101により読み込まれた第1の検索ログデータA10を集計し、集計結果としての第1の集計データB10を出力する。この第1の集計データB10は、図2に示したように、「日付」毎の「キーワード」の利用頻度（回数）を表すデータである。同図に示した例では、2000/11/01に、「カメラ」というキーワードが17回利用されていることがわかる。

【0018】

図1に戻り、第2の集計部103は、読み込部101により読み込まれた第2の検索ログデータA20を集計し、集計結果としての第2の集計データB20を出力する。この第2の集計データB20は、図3に示したように、「日付」毎の「キーワード」の利用頻度（回数）を表すデータである。同図に示した例では、2000/11/01に、「年賀状」というキーワードが3回利用されていることがわかる。

【0019】

ここで、図4を参照して、上述した第1の集計データB10（図2参照）のうちあるキーワード、および第2の集計データB20（図3参照）のうちあるキーワードのそれぞれを時系列で表した場合について説明する。図4において、第1の時系列データD10は、図2に示した第1の集計データB10におけるキーワード「カメラ」に関する日付（12/1日～12/24日）毎の利用頻度の変化

を表している。

【0020】

一方、第2の時系列データD20は、図3に示した第2の集計データB20におけるキーワード「年賀状」に関する日付（12／1日～12／24日）毎の利用頻度の変化を表している。同図からわかるように、インターネット上で利用されるキーワードの利用頻度は、時間的に変化する。また、第1の集計データB10には、キーワードの数分だけの第1の時系列データが存在する。

【0021】

同様にして、第2の集計データB20にも、キーワードの数分だけ第2の時系列データD20が存在する。また、同図に示した例では、第1の時系列データD10における楕円P10のパターンと、第2の時系列データD20における楕円P20のパターンとが近似していることがわかる。

【0022】

図1に戻り、比較部104は、第1の集計データB10に含まれる一つの第1の時系列データと、第2の集計データB20に含まれる一つの第2の時系列データとを、順次、比較し、両者の相関係数を求める。ここで、相関係数が高い場合は、両者の時間的変化（パターン）が近似していることを意味している。一方、相関係数が低い場合には、両者の時間的変化（パターン）が全く異なることを意味している。

【0023】

また、比較部104は、第1の時系列データと第2の時系列データとの相関係数を単純に求めるのではなく、図4に示したように、第1の時系列データD10を基準として、第2の時系列データD20を同図左側へT分だけシフトさせた場合の相関係数を求める。ここで、第2の時系列データD20の起点STが-Tの位置にある場合、第2の時系列データD20と第1の時系列データD10との時差（シフト量）を-Tとする。

【0024】

つぎに、比較部104は、第2の時系列データD20を右へ1シフトさせた状態で、第2の時系列データD20と第1の時系列データD10との相関係数を求

める。以後、比較部104は、第2の時系列データD20を右へ1シフトさせつつ相関係数を順次求める。

【0025】

そして、比較部104は、第2の時系列データD20の起点STが+T（同図では、20日）に位置するまで右シフトさせ、相関係数を求める。この場合、第1の時系列データD10と第2の時系列データD20との時差は、+Tである。すなわち、一つの組み合わせ（第1の時系列データD10および第2の時系列データD20）においては、第2の時系列データD20が-T～+Tの範囲で1シフトされる毎に相関係数が求められる。

【0026】

比較部104は、第1の時系列データD10と第2の時系列データD20との全ての組み合わせに関して、-T～+Tまでの1シフト動作および相関係数の算出を行う。また、図1に示した比較部104は、上述した比較結果として、図5に示した比較結果データCを予測部106へ出力する。

【0027】

この比較結果データCは、「第1の時系列データ」、「第2の時系列データ」、「相関係数」および「時差」から構成されている。「第1の時系列データ」は、第1の集計データB10（図2参照）に含まれるキーワードに関する時系列データである。同図に示した1レコード目の「第1の時系列データ」は、キーワードとしての「カメラ」に関する時系列データである。

【0028】

「第2の時系列データ」は、第2の集計データB20（図3参照）に含まれるキーワードに関する時系列データである。同図に示した1レコード目の「第2の時系列データ」は、キーワードとしての「年賀状」に関する時系列データである。

【0029】

「相関係数」は、「第1の時系列データ」と「第2の時系列データ」との相関係数である。「時差」は、「第1の時系列データ」に対する「第2の時系列データ」のシフト量である。「時差」にマイナス符号が付いている場合には、「第2

の時系列データ」が左シフト（図4参照）されていることを意味している。

【0030】

また、「時差」が0である場合には、シフト量が0（図4参照）であることを意味している。さらに、「時差」に+符号が付いている場合には、「第1の時系列データ」に対して、「第2の時系列データ」が右シフト（図4参照）されていいることを意味している。

【0031】

図1に戻り、入力部105は、キーボード等であり、予測対象のキーワード（以下、予測対象キーワード）および予測日を入力するためのものである。予測部106は、入力部105からの予測対象キーワードおよび予測日をキーとして、比較結果データCに基づいて、予測対象キーワードに関する利用頻度を予測する。

【0032】

図4に示した例では、予測部106は、12月25日以降（楕円Yの部分）の利用頻度を予測する。この予測部106の動作の詳細については、後述する。出力部107は、ディスプレイ、プリンタ等であり、予測部106の予測結果を出力する。

【0033】

つぎに、実施の形態1の動作について、図7～図9に示したフローチャートを参照しつつ説明する。図7に示したステップSA1では、読込部101は、第1の検索ログデータA10（図2参照）および第2の検索ログデータA20（図3参照）を読み込み、これを第1の集計部102および第2の集計部103へ出力する。

【0034】

ステップSA2では、第1の集計部102および第2の集計部103は、集計処理を実行する。すなわち、第1の集計部102は、図2に示した第1の検索ログデータA10を集計し、集計結果としての第1の集計データB10をメモリ（図示略）に格納する。一方、第2の集計部103は、図3に示した第2の検索ログデータA20を集計し、集計結果としての第2の集計データB20をメモリ（

図示略)に格納する。

【0035】

ステップSB3では、比較部104は、比較処理を実行する。すなわち、図8に示したステップSB1では、比較部104は、第1の集計部102のメモリ(図示略)から第1の集計データB10(図2参照)を読み込む。ステップSB2では、比較部104は、第2の集計部103のメモリ(図示略)から第2の集計データB20(図3参照)を読み込む。

【0036】

ステップSB3では、比較部104は、第1の集計データB10(図2参照)から全てのキーワードに関する第1の時系列データを取得したか否かを判断し、この場合、判断結果を「No」とする。ステップSB4では、比較部104は、一つのキーワード(例えば、カメラ)に関する図4に示した第1の時系列データD10を取得する。

【0037】

ステップSB5では、比較部104は、第2の集計データB20(図3参照)から全てのキーワードに関する第2の時系列データを取得したか否かを判断し、この場合、判断結果を「No」とする。ステップSB6では、比較部104は、一つのキーワード(例えば、年賀状)に関する図4に示した第2の時系列データD20を取得する。

【0038】

ステップSB7では、比較部104は、第2の時系列データD20を左にT分だけシフトさせる。ステップSB8では、比較部104は、シフト状態で第2の時系列データD20と第1の時系列データD10との相関係数を以下の(1)式から算出する。

【0039】

【数1】

$$\rho_{xy} = \frac{1}{n} \times \frac{\sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)}{\sigma_x \times \sigma_y} \quad \dots \dots (1)$$

【0040】

上記(1)式において、 ρ_{xy} は、第1の時系列データと第2の時系列データとの相関係数である。nは、第1の時系列データを構成する複数のデータと、第2の時系列データを構成する複数のデータとのうち、時間的に重なるデータの数である。ここで、第1の時系列データに関して、第2の時系列データと時間的に重なる複数のデータは、 $x_1 \sim x_n$ と表現される。一方、第2の時系列データに関して、第1の時系列データと時間的に重なる複数のデータは、 $y_1 \sim y_n$ と表現される。

【0041】

従って、正確には、相関係数 ρ_{xy} は、データ $x_1 \sim x_n$ とデータ $y_1 \sim y_n$ との相関係数である。また、(1)式の μ_x は、データ $x_1 \sim x_n$ の平均値であり、つぎの(2)式で表される。一方、(1)式の μ_y は、データ $y_1 \sim y_n$ の平均値であり、つぎの(3)式で表される。

【0042】

【数2】

$$\mu_x = \frac{\sum_{i=1}^n x_i}{n} \quad \dots \dots (2)$$

【0043】

【数3】

$$\mu_y = \frac{\sum_{i=1}^n y_i}{n} \quad \dots \dots (3)$$

【0044】

また、(1)式の σ_x は、データ $x_1 \sim x_n$ の標準偏差であり、つぎの(4)式で表される。一方、(1)式の σ_y は、データ $y_1 \sim y_n$ の標準偏差であり、つぎの(5)式で表される。

【0045】

【数4】

$$\sigma_x = \sqrt{\frac{n \sum_{i=1}^n (x_i)^2 - (\sum_{i=1}^n x_i)^2}{n(n-1)}} \quad \dots \dots (4)$$

【0046】

【数5】

$$\sigma_y = \sqrt{\frac{n \sum_{i=1}^n (y_i)^2 - (\sum_{i=1}^n y_i)^2}{n(n-1)}} \quad \dots \dots (5)$$

【0047】

図4に示した第2の時系列データD20を左にT分だけシフトさせた状態では、第1の時系列データD10(カメラ)と第2の時系列データD20(年賀状)との相関係数として、例えば、「0.1」が求めらる。比較部104は、比較結果として、図5に示した比較結果データC(第1の時系列データ(カメラ)、第

2の時系列データ（年賀状）、相関係数=0.1、時差=-Tを作成する。

【0048】

ステップSB9では、比較部104は、図4に示した時差0の位置からT分だけ第2の時系列データを右シフトしたか否か、すなわち時差が+Tであるか否かを判断し、この場合、判断結果を「No」とする。ステップSB10では、比較部104は、第2の時系列データ（この場合、第2の時系列データD20：図4参照）を右に1シフトさせる。

【0049】

以後、ステップSB8～ステップSB10が繰り返される。これにより、第1の時系列データD10に対して第2の時系列データD20が右に1シフトされ、相関係数が時差に対応付けられて求められる（図5参照）。すなわち、図2に示した第1の集計データB10における第1の時系列データ（カメラ）と、図3に示した第2の集計データB20における第2の時系列データ（年賀状）とのシフト回数の組み合わせに関する相関係数が、時差に対応付けられて求められる。

【0050】

ここで、第2の時系列データD20（年賀状）を右へ9シフトさせた場合の第1の時系列データD10（カメラ）を図6に示す。同図からわかるように、楕円P10のパターンと、楕円P20のパターンとが時間的にほぼ一致している。この場合の相関係数は、例えば、0.68であるものとし、非常に高い値である。

【0051】

そして、ステップSB9の判断結果が「Yes」になると、ステップSB5では、比較部104は、第2の集計データB20（図3参照）から全てのキーワードに関する第2の時系列データを取得したか否かを判断し、この場合、判断結果を「No」とする。ステップSB6では、比較部104は、つぎのキーワード（例えば、ラジオ）に関する第2の時系列データ（図示略）を取得する。

【0052】

ステップSB7では、比較部104は、第2の時系列データ（ラジオ）を左にT分だけシフトさせる。ステップSB8では、比較部104は、シフト状態で第2の時系列データ（ラジオ）と第1の時系列データD10（カメラ）との相関係

数を前述した(1)式から算出する。以後、前述した動作が繰り返される。これにより、第1の時系列データD10(カメラ)に対して第2の時系列データ(ラジオ)が右に1シフトされ、相関係数が時差に対応付けられて求められる。

【0053】

そして、ステップSB9の判断結果が「Yes」になると、比較部104は、ステップSB5の判断結果を「No」として、以後、ステップSB3～ステップSB10を繰り返す。

【0054】

そして、全てのキーワードに関する第1の時系列データと、全てのキーワードに関する第2の時系列データとの組み合わせであって、さらにシフト回数の組み合わせの相関係数が求められると、ステップSB3の判断結果が「Yes」となる。ステップSB11では、比較部104は、図5に示した比較結果データCを予測部106へ出力する。

【0055】

図7に示したステップSA4では、予測部106は、比較結果データCに基づいて予測処理を実行する。具体的には、図9に示したステップSC1では、ユーザは、入力部105を用いて、予測対象キーワードおよび日付を入力する。

【0056】

この場合、ユーザは、図4に示した橜円Y部分の利用頻度(キーワード:カメラ)を予測させるために、ユーザは、予測対象キーワードとして「カメラ」、日付として「12/25」を入力する。これにより、第1の時系列データD10(カメラ)に関して、12/25以降の利用頻度が予測される。

【0057】

ステップSC2では、予測部106は、ステップSC1で入力された予測対象キーワードK(カメラ)と、日付D(12/25)とを読み込む。ステップSC3では、予測部106は、図5に示した比較結果データCから、予測対象キーワードK(カメラ)をキーとして、相関係数が最も高いキーワードを基準キーワードKWとして取得する。

【0058】

この場合、図5および図6に示した第1の時系列データ（カメラ）、第2の時系列データ（年賀状）、相関係数=0.68、時差=+9の組み合わせが最も相関係数が高いものとする。この組み合わせは、図6に示したグラフである。

【0059】

従って、この組み合わせでは、予測対象キーワードKが「カメラ」、基準キーワードKWが「年賀状」となり、基準キーワードKW（年賀状）に対応する図6に示した第2の時系列データD20に基づいて、12/25日以降の予測対象キーワードK（カメラ）に関する利用頻度が予測される。

【0060】

具体的には、ステップSC4では、予測部106は、図4に示した第2の時系列データD20を、図6に示したように時差9分だけ右ヘシフトさせる。さらに、予測部106は、第1の時系列データD10とほぼ一致するように、図6に示した第2の時系列データD20を上方ヘシフトさせる。これにより、図10に示した状態となる。

【0061】

つぎに、予測部106は、図10に示した第2の時系列データD20における12/25以降のデータ（楕円Y）を用いて、第1の時系列データD10における12/25以降の利用頻度を予測する。具体的には、予測部106は、最小二乗法を用いた回帰分析により、予測部分の直線を表すつぎの(6)式を求める。

【0062】

【数6】

$$p = md + k \quad \dots\dots (6)$$

【0063】

(6)式において、pは、直線の関数であり、予測対象キーワードKの利用頻度に関する予測値である。mは、直線の傾きであり、つぎの(7)式で表される。dは、基準キーワードKW（年賀状）に対応する第2の時系列データD20の値である。kは、直線におけるy切片であり、つぎの(8)式で表される。

【0064】

【数7】

$$m = \frac{n \left(\sum_{i=1}^n (x_i \times y_i) \right) \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n y_i \right)}{n \left(\sum_{i=1}^n (x_i^2) \right) - \left(\sum_{i=1}^n x_i \right)^2} \quad \dots \dots (7)$$

【0065】

【数8】

$$k = \frac{\left(\sum_{i=1}^n y_i \right) \left(\sum_{i=1}^n (x_i^2) \right) - \left(\sum_{i=1}^n x_i \right) \left(\sum_{i=1}^n x_i y_i \right)}{n \left(\sum_{i=1}^n (x_i^2) \right) - \left(\sum_{i=1}^n x_i \right)^2} \quad \dots \dots (8)$$

【0066】

(7) 式および(8)式において、nは、第1の時系列データを構成する複数のデータと、第2の時系列データを構成する複数のデータとのうち、時間的に重なるデータの数である。第1の時系列データに関して、第2の時系列データと時間的に重なる複数のデータは、 $x_1 \sim x_n$ と表現される。一方、第2の時系列データに関して、第1の時系列データと時間的に重なる複数のデータは、 $y_1 \sim y_n$ と表現される。

【0067】

そして、上記利用頻度の予測を終えると、図7に示したステップSA5では、予測部106は、ステップSA4で求められた予測利用頻度データを出力部107から出力する。この予測利用頻度データは、図10に示した橢円Yの部分に相当し、第1の時系列データD10(カメラ)に関する12/25日以降の利用頻度である。

【0068】

以上説明したように、実施の形態1によれば、図4に示した第1の時系列データD10に対応する第1のパターンと、第2の時系列データD20に対応する第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行い、所定時間単位毎（1シフト毎）に、第1の時系列データD10と第2の時系列データD20との相関係数を算出し、最も高い値の相関係数に対応する第1の時系列データD10と第2の時系列データD20との組に対応する第2のパターンに基づいて、組を構成する第1の時系列データD10の利用頻度を予測（図10参照）しているため、従来のように一つの時系列情報に基づいて、予測を行う場合に比して、情報の利用頻度を高精度で予測することができる。

【0069】

また、実施の形態1によれば、第1のグループに属する複数の第1の集計データB10における複数の第1の時系列情報と、第2の集計データB20における複数の第2の時系列情報との全ての組み合わせに関して、時間的な操作を行うようにしたので、組み合わせの増加により、相関係数が高い組み合わせが存在する確率が高まり、情報の利用頻度を高精度で予測することができる。

【0070】

また、実施の形態1によれば、図4に示したように、第1の時系列データD10を基準として、第2の時系列データD20を時間軸上で、順次、所定時間単位でシフトさせるようにしたので、流行等の時期がずれている場合であっても、実際に、相関係数が高い組も特定されるため、情報の利用頻度を高精度で予測することができる。

【0071】

また、実施の形態1によれば、第1の検索ログデータA10と第2の検索ログデータA20とが異なる収集経路を介して収集されるため、幅広い分野からの情報に基づいて利用頻度の予測を行うことができ、予測の精度を高めることができる。

【0072】

また、実施の形態1によれば、第1の時系列データD10（カメラ）および第

2の時系列データD20(年賀状)を、インターネット上のキーワード検索エンジンにおけるキーワードの利用頻度の時系列情報としたので、従来のように一つの時系列情報に基づいて、予測を行う場合に比して、インターネットにおけるキーワードの利用頻度を高精度で予測することができる。

【0073】

図11は、実施の形態1の効果を説明する図である。この図において、第1の時系列データD10は、前述したキーワードとしての「カメラ」に対応している。この第1の時系列データD10において、12/24以前のデータは、実測値である。また、第1の時系列データD10において、12/25以降のデータは、前述した手法により予測されたデータである。

【0074】

また、12/25以降において、実測データJは、キーワードとしての「カメラ」の利用頻度に関する実測データである。この図からわかるように、実測データJと、12/25以降の第1の時系列データD10(予測値)とがほぼ一致しており、平均誤差が17%である。これに対して、従来の時系列分析による予測では、実測値と予測値との平均誤差が40%となる。このように、実施の形態1の手法では、従来に比べて、23%も平均誤差が少なくなっており、予測精度が飛躍的に高くなる。

【0075】

(実施の形態2)

さて、前述した実施の形態1では、図1に示した予測部106で予測された予測利用頻度データにソートをかけるようにしてもよい。以下では、この場合を実施の形態2として説明する。

【0076】

図12は、本発明にかかる実施の形態2の構成を示すブロック図である。この図において、図1の各部に対応する部分には同一の符号を付ける。同図に示した情報利用頻度予測装置200においては、ソート部201および取扱商品データベース202が新たに設けられている。

【0077】

ソート部201は、予測部106からの予測利用頻度データに基づいて、キーワードに対応する取扱商品と、該キーワードの予測利用頻度との関係を表す予測利用頻度データE1（図14参照）を作成する機能を備えている。また、ソート部201は、図14に示した予測利用頻度データE1の予測利用頻度をソートキーとして、例えば、昇順にソートをかけ、ソート済み予測利用頻度データE2を作成する機能を備えている。

【0078】

図12に戻り、取扱商品データベース202は、上記キーワードと、店舗での取扱商品との対応関係を表すデータベースである。この取扱商品データベース202は、ソート部201で予測利用頻度データE1を作成する際に利用される。

【0079】

つぎに、実施の形態2の動作について、図13に示したフローチャートを参照しつつ説明する。なお、図13に示したステップSD1～ステップSD4は、図7に示したステップSA1～ステップSA4と同様の処理であるため、その説明を省略する。

【0080】

この場合、予測部106から複数のキーワードに関する予測利用頻度データが出力されると、図13に示したステップSD5では、ソート部201は、取扱商品データベース202を参照して、キーワードを取扱商品に変換し、図14に示した予測利用頻度データE1を作成する。

【0081】

つぎに、ソート部201は、予測利用頻度をソートキーとして、予測利用頻度データE1を昇順にソートする。このソート結果は、ソート済み予測利用頻度データE2である。ステップSD6では、ソート部201は、ソート済み予測利用頻度データE2を出力部107より出力する。

【0082】

以上説明したように、実施の形態2によれば、図14に示したように、複数の予測結果に対して、利用頻度をキーとしてソートをかけるようにしたので、予測された利用頻度のソート結果から、例えば、電子商取引を行っている店舗等にお

ける売れ筋商品の情報を容易に得ることができ、販売促進を図ることができる。

【0083】

(実施の形態3)

さて、前述した実施の形態1では、図4に示したように、第1の時系列データD10に対して、パターンを変えることなく単純に第2の時系列データD20を左から右へ1シフトさせ、両者の相関係数を求めた例について説明したが、このようなシフト処理に加えて、第2の時系列データD20のパターンの時間軸を伸縮させ、両者の相関係数を求めるようにしてもよい。以下では、この場合を実施の形態3として説明する。

【0084】

図15は、本発明にかかる実施の形態3の構成を示すブロック図である。この図において、図1の各部に対応する部分には同一の符号を付ける。同図に示した情報利用頻度予測装置300においては、図1に示した比較部104に代えて、比較部301が設けられている。

【0085】

比較部301は、第1の集計データB10に含まれる一つの第1の時系列データと、第2の集計データB20に含まれる一つの第2の時系列データとを、順次、比較し、比較結果データF（図16参照）を出力する。

【0086】

具体的には、比較部301は、図17に示したように、第1の時系列データD100を基準として、第2の時系列データD200を同図左側へT分だけシフト（図示略）させる。つぎに、比較部301は、シフトされた状態の第2の時系列データD200のパターンの時間軸を、例えば、伸縮率d=50%（図18参照）で縮小する。この伸縮率d=50%は、予め決められた縮小率のしきい値である。

【0087】

すなわち、比較部301は、第2の時系列データD200を左側へT分（最大値）シフトさせた後、さらに、時間軸を50%で縮小する。この状態で、比較部301は、前述した（1）式を用いて第1の時系列データD100と第2の時系

列データD200（左シフト+縮小）との相関係数を求める。

【0088】

つぎに、比較部301は、第2の時系列データD200の時間軸を、単位伸張率 Δd （例えば、10%：図18参照）で伸張させた状態で、第2の時系列データD200と第1の時系列データD100との相関係数を求める。以後、比較部301は、伸縮率dが伸張率z（例えば、200%：図18参照）になるまで、第2の時系列データD200の時間軸を単位伸張率 Δd 分だけ伸張しつつ相関係数を順次求める。

【0089】

すなわち、第2の時系列データD200は、シフト位置が固定された状態で、伸縮率 $d = 50\% \sim 200\%$ の範囲で単位伸張率 Δd 刻みで徐々に伸張されるのである。

【0090】

そして、伸縮率dが伸張率z（=200%）になると、比較部301は、第2の時系列データD200を右へ1シフトさせた後、シフトされた状態の第2の時系列データD200のパターンの時間軸を、伸縮率 $d = 50\%$ で縮小する。この状態で、比較部301は、第2の時系列データD200と第1の時系列データD100との相関係数を求める。

【0091】

以後、比較部301は、シフト位置を固定した状態で、伸縮率 $d = 50\% \sim 200\%$ の範囲で単位伸張率 Δd 刻みで徐々に第2の時系列データD200を伸張する毎に相関係数を求める。第2の時系列データD200のシフトおよび伸張は、第2の時系列データD200が右側にT分だけシフトされるまで行われる。

【0092】

上述したように、比較部301は、第1の時系列データD100と第2の時系列データD200との全ての組み合わせに関して、 $-T \sim +T$ までの1シフト動作、時間軸の伸縮および相関係数の算出を行う。

【0093】

図16は、比較部301から出力される比較結果データFを示す図である。こ

の比較結果データFは、「第1の時系列データ」、「第2の時系列データ」、「相関係数」、「時差」および「伸縮率」から構成されている。これらの「第1の時系列データ」、「第2の時系列データ」および「時差」は、図5に示した比較結果データCにおける「第1の時系列データ」、「第2の時系列データ」および「時差」と同義である。

【0094】

すなわち、「第1の時系列データ」は、第1の集計データB10（図2参照）に含まれるキーワードに関する時系列データである。「第2の時系列データ」は、第2の集計データB20（図3参照）に含まれるキーワードに関する時系列データである。「時差」は、「第1の時系列データ」に対する「第2の時系列データ」のシフト量である。

【0095】

「相関係数」は、第2の時系列データがシフトおよび伸張された状態における「第1の時系列データ」と「第2の時系列データ」との相関係数である。「伸縮率」は、第2の時系列データにおける時間軸の伸張率であり、前述したように単位伸張率 Δd （=10%）刻みで50%～200%までの範囲をとる。

【0096】

つぎに、実施の形態3の動作について、図20～図22に示したフローチャートを参照しつつ説明する。なお、図20に示したステップSE1、ステップSE2、ステップSE4およびステップSE5は、図7に示したステップSA1、ステップSA2、ステップSA4およびステップSA5に対応している。

【0097】

また、図21に示したステップSF1～ステップSF7およびステップSF9～ステップSF11は、図8に示したステップSB1～ステップSB7、ステップSB9～ステップSB11に対応している。

【0098】

図20に示したステップSE1では、読み込部101は、第1の検索ログデータA10（図2参照）および第2の検索ログデータA20（図3参照）を読み込み、これらを第1の集計部102および第2の集計部103へ出力する。

【0099】

ステップSE2では、第1の集計部102および第2の集計部103は、集計処理を実行し、第1の集計データB10（図2参照）および第2の集計データB20（図3参照）をそれぞれのメモリ（図示略）に格納する。

【0100】

ステップSE3では、比較部301は、比較処理を実行する。すなわち、図21に示したステップSF1では、第1の集計データB10（図2参照）が読み込まれ、ステップSF2では、第2の集計データB20（図3参照）が読み込まれる。

【0101】

ステップSF3では、比較部301は、ステップSB3（図8参照）と同様にして、第1の集計データB10（図2参照）から全てのキーワードに関する第1の時系列データを取得したか否かを判断し、この場合、判断結果を「No」とする。ステップSF4では、比較部301は、一つのキーワード（例えば、カメラ）に関する図17に示した第1の時系列データD100を取得する。

【0102】

ステップSF5では、比較部301は、ステップSB5（図8参照）と同様にして、第2の集計データB20（図3参照）から全てのキーワードに関する第2の時系列データを取得したか否かを判断し、この場合、判断結果を「No」とする。

【0103】

ステップSF6では、比較部301は、一つのキーワード（例えば、年賀状）に関する図17に示した第2の時系列データD200を取得する。ここで、第2の時系列データD200のグラフ波形は、第1の時系列データD100のパターンが時間軸上で縮小された波形に近似していることがわかる。

【0104】

ステップSF7では、比較部301は、ステップSB7（図8参照）と同様にして、第2の時系列データD200を左にT分だけシフトさせる。ステップSF8では、比較部301は、第2の時系列データD200を伸縮させる伸縮処理を

実行する。

【0105】

具体的には、図22に示したステップSG1では、比較部301は、伸縮率dを50%として、第2の時系列データD200の時間軸を縮小させる。ステップSG2では、比較部301は、シフト状態および伸縮状態で第2の時系列データD200と第1の時系列データD100との相関係数（例えば、0.1：図16参照）を前述した(1)式から算出する。

【0106】

つぎに、比較部301は、比較結果として、図16に示した比較結果データF（第1の時系列データ（カメラ）、第2の時系列データ（年賀状）、相関係数=0.1、時差=-T、伸縮率=50%）を作成する。

【0107】

ステップSG3では、比較部301は、伸縮率d (=50%)に単位伸張率△d (=10%)を加算して、伸縮率dを60%にする。つまり、比較部301は、第2の時系列データD200の時間軸の伸縮率dを60%にし、10%分だけ伸張させる。

【0108】

ステップSG4では、比較部301は、伸縮率d (=60%)が伸張率z (=200%)であるか否かを判断し、この場合、判断結果を「No」とする。以後、伸縮率dが200%になるまで、ステップSG2～ステップSG4が繰り返される。

【0109】

以後、比較部301は、第2の時系列データD200のシフト位置を固定した状態で、伸縮率d = 60%～200%の範囲で単位伸張率△d刻みで徐々に第2の時系列データD200を伸張する毎、相関係数を順次求める。

【0110】

そして、ステップSG4の判断結果が「Yes」になると、図21に示したステップSF9では、比較部301は、ステップSB9（図8参照）と同様にして、第2の時系列データD200をT分だけ右シフトしたか否かを判断し、この場

合、判断結果を「No」とする。ステップSF10では、比較部301は、第2の時系列データ（この場合、第2の時系列データD200：図17参照）を右に1シフトさせる。

【0111】

以後、ステップSF8～ステップSF10が繰り返される。これにより、第1の時系列データD100に対して第2の時系列データD200が右に1シフトされるとともに、時間軸が50%～200%の範囲で伸縮され、相関係数が時差および伸縮率に対応付けられて求められる（図16参照）。

【0112】

すなわち、図2に示した第1の集計データB10における第1の時系列データ（カメラ）と、図3に示した第2の集計データB20における第2の時系列データ（年賀状）とのシフト回数および伸縮回数の組み合わせに関する相関係数が、時差および伸縮率に対応付けられて求められる。

【0113】

ここで、第2の時系列データD200（年賀状）を右へある値でシフトさせた状態で、時間軸をある値で伸縮させた場合の第1の時系列データD100（カメラ）を図19に示す。同図からわかるように、楕円P100のパターンと、楕円P200のパターンとが時間的にほぼ一致している。この場合の相関係数は、例えば、0.68であるものとし、非常に高い値である。

【0114】

そして、ステップSF9の判断結果が「Yes」になると、ステップSF5では、比較部301は、第2の集計データB20（図3参照）から全てのキーワードに関する第2の時系列データを取得したか否かを判断し、この場合、判断結果を「No」とする。ステップSF6では、比較部301は、つぎのキーワード（例えば、ラジオ）に関する第2の時系列データ（図示略）を取得する。

【0115】

ステップSF7では、比較部301は、第2の時系列データ（ラジオ）を左にT分だけシフトさせる。ステップSF8では、比較部301は、シフト状態で第2の時系列データ（ラジオ）と第1の時系列データD100（カメラ）との相関

係数を前述した(1)式から算出する。

【0116】

以後、前述した動作が繰り返される。そして、全てのキーワードに関する第1の時系列データと、全てのキーワードに関する第2の時系列データとの組み合わせであって、さらにシフト回数および伸縮回数の組み合わせの相関係数が求められると、ステップSF3の判断結果が「Yes」となる。ステップSF11では、比較部301は、図16に示した比較結果データFを予測部106へ出力する。

【0117】

図20に示したステップSE4では、予測部106は、図9に示した予測処理と同様にして、比較結果データF(図16参照)に基づいて予測処理を実行する。図19に示した第1の時系列データD100と、第2の時系列データD200との組み合わせの相関係数が最も高い場合には、予測部106は、前述した手法により、第2の時系列データD200から第1の時系列データD100におけるキーワード(カメラ)の利用頻度を予測する。

【0118】

以上説明したように、実施の形態3によれば、図17～図19に示した第1の時系列データD100を基準として、第2の時系列データD200を時間軸上で、順次、所定時間単位でシフトさせ、シフトされた第2の時系列データD200を時間的に、所定伸縮単位(単位伸張率 Δd)で、順次、伸縮させるようにしたので、流行等の時期のずれている場合や、流行等のスピードが相違している場合であっても、実際上、相関係数が高い組も特定されるため、情報の利用頻度をさらに高精度で予測することができる。

【0119】

また、実施の形態3によれば、図17～図19に示した第1の時系列データD100を基準として、第2の時系列データD200を時間的に、所定伸縮単位(単位伸張率 Δd)で、順次、伸縮させるようにしたので、流行等のスピードが相違している場合であっても、実際上、相関係数が高い組も特定されるため、情報の利用頻度をさらに高精度で予測することができる。

【0120】

以上本発明にかかる実施の形態1～3について図面を参照して詳述してきたが、具体的な構成例はこれらの実施の形態1～3に限られるものではなく、本発明の要旨を逸脱しない範囲の設計変更等があっても本発明に含まれる。

【0121】

例えば、前述した実施の形態1～3においては、二つの検索エンジンのそれぞれの検索ログデータを第1の検索ログデータA10および第2の検索ログデータA20（図1、図12、図15）とし、二つの第1の集計部102および第2の集計部103でそれぞれ集計する例について説明したが、これに限られることなく、一つの集計部とする構成としてもよい。

【0122】

この場合には、上記集計部は、一つの検索エンジンの検索ログデータを二つのグループに分け、それぞれのグループに関して集計を行い、第1の集計データB10および第2の集計データB20を出力すればよい。この場合には、集計部の構成が一つで済むため、構成を簡素化することができる。

【0123】

また、前述した実施の形態1～3においては、利用頻度の予測対象としてインターネット検索におけるキーワードを一例にとって説明したが、キーワードに限られることなく、如何なる情報であってよい。

【0124】

また、前述した実施の形態1～3においては、前述した情報利用頻度予測装置100（図1参照）、情報利用頻度予測装置200（図12参照）または情報利用頻度予測装置300（図15参照）の機能を実現するための情報利用頻度予測プログラムを図23に示したコンピュータ読み取り可能な記録媒体500に記録して、この記録媒体500に記録された情報利用頻度予測プログラムを同図に示したコンピュータ400に読み込ませ、実行することにより、キーワード（情報）に関する一連の利用頻度予測処理を実行するようにしてもよい。

【0125】

コンピュータ400は、上記情報利用頻度予測プログラムを実行するCPU4

10と、キーボード、マウス等の入力装置420と、各種データを記憶するROM (Read Only Memory) 430と、演算パラメータ等を記憶するRAM (Random Access Memory) 440と、記録媒体500から情報利用頻度予測プログラムを読み取る読み取り装置450と、ディスプレイ、プリンタ等の出力装置460と、各部を接続するバス470とから構成されている。

【0126】

CPU410は、読み取り装置450を経由して記録媒体500に記録されている情報利用頻度予測プログラムを読み込んだ後、情報利用頻度予測プログラムを実行することにより、前述したキーワード（情報）に関する一連の利用頻度予測処理を実行する。なお、記録媒体500には、光ディスク、フロッピーディスク、ハードディスク等の可搬型の記録媒体が含まれることはもとより、ネットワークのようにデータを一時的に記録保持するような伝送媒体も含まれる。

【0127】

（付記1）コンピュータを、

第1の時系列情報の利用頻度に関する時間的変化を表す第1のパターンと、第2の時系列情報の利用頻度に関する時間的変化を表す第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行う時間的操作手段、

前記所定時間単位毎に、前記第1の時系列情報と前記第2の時系列情報との相関係数を算出する相関係数算出手段、

前記相関係数算出手段により算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組を特定する組特定手段、

前記組に対応する第2のパターンに基づいて、前記組を構成する第1の時系列情報の利用頻度を予測する予測手段、

として機能させることを特徴とする情報利用頻度予測プログラム。

（付記2）前記時間的操作手段は、第1のグループに属する複数の第1の時系列情報と、第2のグループに属する複数の第2の時系列情報との全ての組み合わせに関して、前記時間的な操作を行うことを特徴とする付記1に記載の情報利用頻度予測プログラム。

(付記3) コンピュータを、前記予測手段における複数の予測結果に対して、前記利用頻度をキーとしてソートをかけるソート手段として機能させることを特徴とする付記2に記載の情報利用頻度予測プログラム。

(付記4) 前記時間的操作手段は、前記第1の時系列情報を基準として、前記第2の時系列情報を時間軸上で、順次、所定時間単位でシフトさせることを特徴とする付記1～3のいずれか一つに記載の情報利用頻度予測プログラム。

(付記5) 前記時間的操作手段は、前記第1の時系列情報を基準として、前記第2の時系列情報を時間的に、所定伸縮単位で、順次、伸縮させることを特徴とする付記1～3のいずれか一つに記載の情報利用頻度予測プログラム。

(付記6) 前記時間的操作手段は、前記第1の時系列情報を基準として、前記第2の時系列情報を時間軸上で、順次、所定時間単位でシフトさせ、シフトされた前記第2の時系列情報を時間的に、所定伸縮単位で、順次、伸縮させることを特徴とする付記1～3のいずれか一つに記載の情報利用頻度予測プログラム。

(付記7) 前記第1の時系列情報および前記第2の時系列情報は、インターネット上のキーワード検索エンジンにおけるキーワードの利用頻度の時系列情報であることを特徴とする付記1～6のいずれか一つに記載の情報利用頻度予測プログラム。

(付記8) 前記第1の時系列情報と前記第2の時系列情報とは、異なる収集経路を介して収集されることを特徴とする付記1～7のいずれか一つに記載の情報利用頻度予測プログラム。

(付記9) 前記第1の時系列情報と前記第2の時系列情報とは、同一の収集経路を介して収集され、収集された時系列情報が二つにグループ化されたものであることを特徴とする付記1～7のいずれか一つに記載の情報利用頻度予測プログラム。

(付記10) 第1の時系列情報の利用頻度に関する時間的变化を表す第1のパターンと、第2の時系列情報の利用頻度に関する時間的变化を表す第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行う時間的操作手段と、

前記所定時間単位毎に、前記第1の時系列情報と前記第2の時系列情報との相

関係数を算出する相関係数算出手段と、

前記相関係数算出手段により算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組を特定する組特定手段と、

前記組に対応する第2のパターンに基づいて、前記組を構成する第1の時系列情報の利用頻度を予測する予測手段と、

を備えたことを特徴とする情報利用頻度予測装置。

(付記11) 前記時間的操手段は、前記第1の時系列情報を基準として、前記第2の時系列情報を時間軸上で、順次、所定時間単位でシフトさせることを特徴とする付記10に記載の情報利用頻度予測装置。

(付記12) 前記時間的操手段は、前記第1の時系列情報を基準として、前記第2の時系列情報を時間的に、所定伸縮単位で、順次、伸縮させることを特徴とする付記10または付記11に記載の情報利用頻度予測装置。

(付記13) 第1の時系列情報の利用頻度に関する時間的变化を表す第1のパターンと、第2の時系列情報の利用頻度に関する時間的变化を表す第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行う時間的操作工程と、

前記所定時間単位毎に、前記第1の時系列情報と前記第2の時系列情報との相関係数を算出する相関係数算出工程と、

前記相関係数算出工程で算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組を特定する組特定工程と、

前記組に対応する第2のパターンに基づいて、前記組を構成する第1の時系列情報の利用頻度を予測する予測工程と、

を含むことを特徴とする情報利用頻度予測方法。

(付記14) 前記時間的操手段では、前記第1の時系列情報を基準として、前記第2の時系列情報を時間軸上で、順次、所定時間単位でシフトさせることを特徴とする付記13に記載の情報利用頻度予測方法。

(付記15) 前記時間的操手段では、前記第1の時系列情報を基準として、前

記第2の時系列情報を時間的に、所定伸縮単位で、順次、伸縮させることを特徴とする付記13または付記14に記載の情報利用頻度予測方法。

【0128】

【発明の効果】

以上説明したように、本発明によれば、第1の時系列情報に対応する第1のパターンと、第2の時系列情報に対応する第2のパターンとの相対関係に対して、順次、所定時間単位で時間的な操作を行い、所定時間単位毎に、第1の時系列情報と第2の時系列情報との相関係数を算出し、最も高い値の相関係数に対応する第1の時系列情報と第2の時系列情報との組に対応する第2のパターンに基づいて、組を構成する第1の時系列情報の利用頻度を予測しているため、従来のように一つの時系列情報に基づいて、予測を行う場合に比して、情報の利用頻度を高精度で予測することができるという効果を奏する。

【0129】

また、本発明によれば、第1のグループに属する複数の第1の時系列情報と、第2のグループに属する複数の第2の時系列情報との全ての組み合わせに関して、時間的な操作を行うようにしたので、組み合わせの増加により、相関係数が高い組み合わせが存在する確率が高まり、情報の利用頻度をさらに高精度で予測することができるという効果を奏する。

【0130】

また、本発明によれば、複数の予測結果に対して、利用頻度をキーとしてソートをかけるようにしたので、予測された利用頻度のソート結果から、例えば、電子商取引を行っている店舗等における売れ筋商品の情報を容易に得ることができ、販売促進を図ることができるという効果を奏する。

【0131】

また、本発明によれば、第1の時系列情報を基準として、第2の時系列情報を時間軸上で、順次、所定時間単位でシフトさせるようにしたので、流行等の時期がズれている場合であっても、実際上、相関係数が高い組も特定されるため、情報の利用頻度をさらに高精度で予測することができるという効果を奏する。

【0132】

また、本発明によれば、第1の時系列情報を基準として、第2の時系列情報を時間的に、所定伸縮単位で、順次、伸縮させるようにしたので、流行等のスピードが相違している場合であっても、実際上、相関係数が高い組も特定されるため、情報の利用頻度をさらに高精度で予測することができるという効果を奏する。

【0133】

また、本発明によれば、第1の時系列情報を基準として、第2の時系列情報を時間軸上で、順次、所定時間単位でシフトさせ、シフトされた第2の時系列情報を時間的に、所定伸縮単位で、順次、伸縮させるようにしたので、流行等の時期のずれている場合や、流行等のスピードが相違している場合であっても、実際上、相関係数が高い組も特定されるため、情報の利用頻度をさらに高精度で予測することができるという効果を奏する。

【0134】

また、本発明によれば、第1の時系列情報および第2の時系列情報を、インターネット上のキーワード検索エンジンにおけるキーワードの利用頻度の時系列情報としたので、従来のように一つの時系列情報に基づいて、予測を行う場合に比して、インターネットにおけるキーワードの利用頻度を高精度で予測することができるという効果を奏する。

【0135】

また、本発明によれば、第1の時系列情報と第2の時系列情報とが異なる収集経路を介して収集されるため、幅広い分野からの情報に基づいて利用頻度の予測を行うことができ、さらに予測の精度を高めることができるという効果を奏する。

【図面の簡単な説明】

【図1】

本発明にかかる実施の形態1の構成を示すブロック図である。

【図2】

図1に示した第1の検索ログデータA10および第1の集計データB10の一例を示す図である。

【図3】

図1に示した第2の検索ログデータA20および第2の集計データB20の一例を示す図である。

【図4】

同実施の形態1における第1の時系列データD10および第2の時系列データD20の一例を示す図である。

【図5】

図1に示した比較結果データCの一例を示す図である。

【図6】

同実施の形態1における右シフトを説明する図である。

【図7】

同実施の形態1の動作を説明するフローチャートである。

【図8】

図7に示した比較処理を説明するフローチャートである。

【図9】

図7に示した予測処理を説明するフローチャートである。

【図10】

同実施の形態1の動作を説明する図である。

【図11】

同実施の形態1の効果を説明する図である。

【図12】

本発明にかかる実施の形態2の構成を示すブロック図である。

【図13】

同実施の形態2の動作を説明するフローチャートである。

【図14】

図13に示したソート処理を説明する図である。

【図15】

本発明にかかる実施の形態3の構成を示すブロック図である。

【図16】

図15に示した比較結果データFの一例を示す図である。

【図17】

同実施の形態3の動作を説明する図である。

【図18】

同実施の形態3の動作を説明する図である。

【図19】

同実施の形態3の動作を説明する図である。

【図20】

同実施の形態3の動作を説明するフローチャートである。

【図21】

図20に示した比較処理を説明するフローチャートである。

【図22】

図21に示した伸縮処理を説明するフローチャートである。

【図23】

実施の形態1～3の変形例の構成を示すブロック図である。

【符号の説明】

100 情報利用頻度予測装置

102 第1の集計部

103 第2の集計部

104 比較部

105 入力部

106 予測部

200 情報利用頻度予測装置

201 ソート部

300 情報利用頻度予測装置

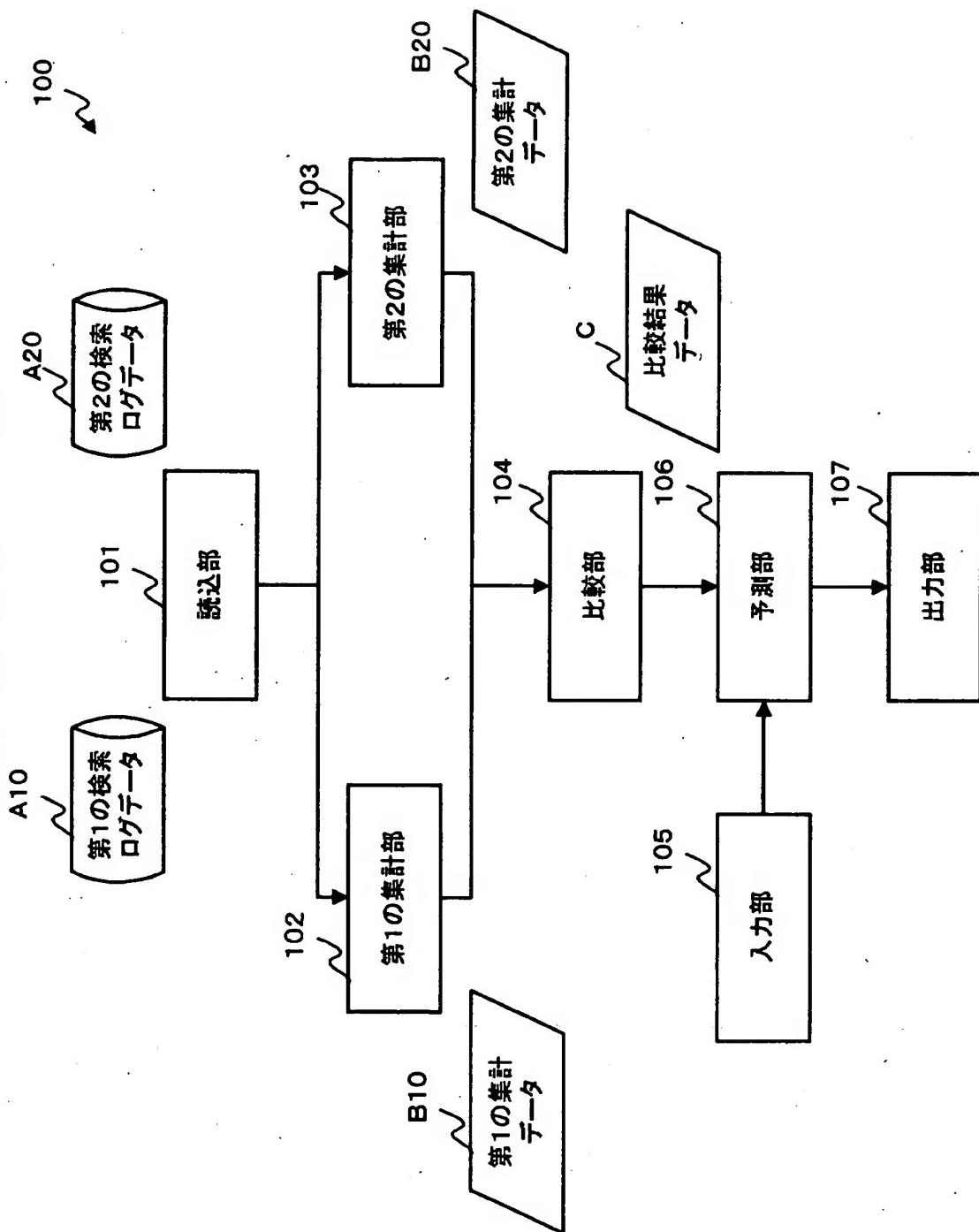
301 比較部

【書類名】

図面

【図1】

実施の形態1の構成を示すブロック図



【図2】

図1に示した第1の検索ログデータA10および第1の集計データB10の一例を示す図

A10

日付	時刻	キーワード
2000/11/01	01:10:00	カメラ
2000/11/01	01:10:25	フィルム
2000/11/01	01:10:26	カメラ
2000/11/01	01:10:30	フィルム
2000/11/01	01:10:20	カメラ

B10

日付	キーワード	利用頻度
2000/11/01	カメラ	17
2000/11/01	フィルム	10
2000/11/01	ラッシュ	10
2000/11/01	三脚	12

【図3】

図1に示した第2の検索ログデータA20および第2の集計データB20の一例を示す図

The diagram illustrates two tables, A20 and B20, representing search logs and statistics.

A20 (Search Log Data):

日付	時刻	キーワード
2000/11/01	11:10:50	年賀状
2000/11/01	11:10:41	年賀状
2000/11/01	11:10:56	スタンプ
2000/11/01	11:11:10	年賀状
2000/11/01	11:11:35	スタンプ

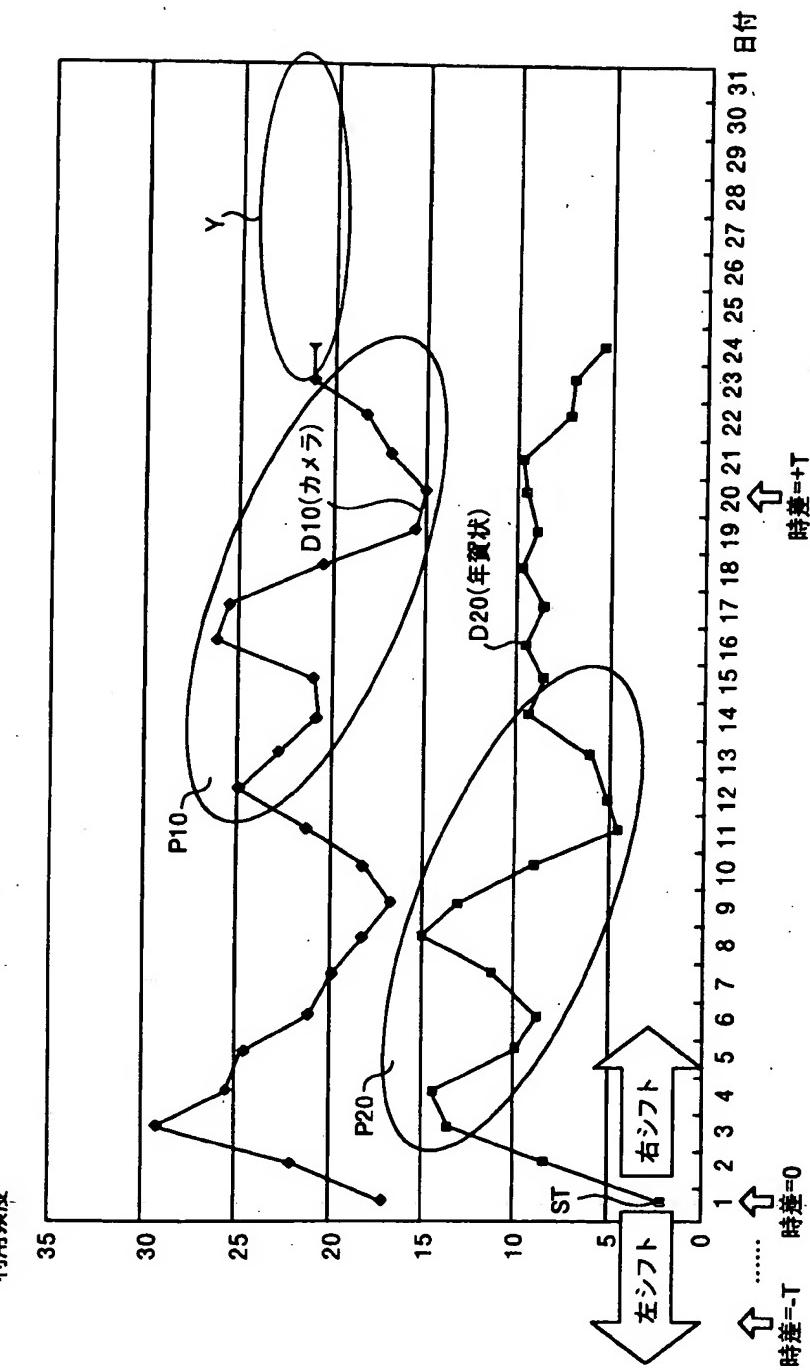
B20 (Statistics Data):

日付	キーワード	利用頻度
2000/11/01	年賀状	3
2000/11/01	ラジオ	13
2000/11/01	スタンプ	4
2000/11/01	電話	11

An arrow points from the text "図1に示した第2の検索ログデータA20および第2の集計データB20の一例を示す図" to the A20 table. Another arrow points from the text "図1に示した第2の検索ログデータA20および第2の集計データB20の一例を示す図" to the B20 table.

【図4】

実施の形態1における第1の時系列データD10および
第2の時系列データD20の一例を示す図



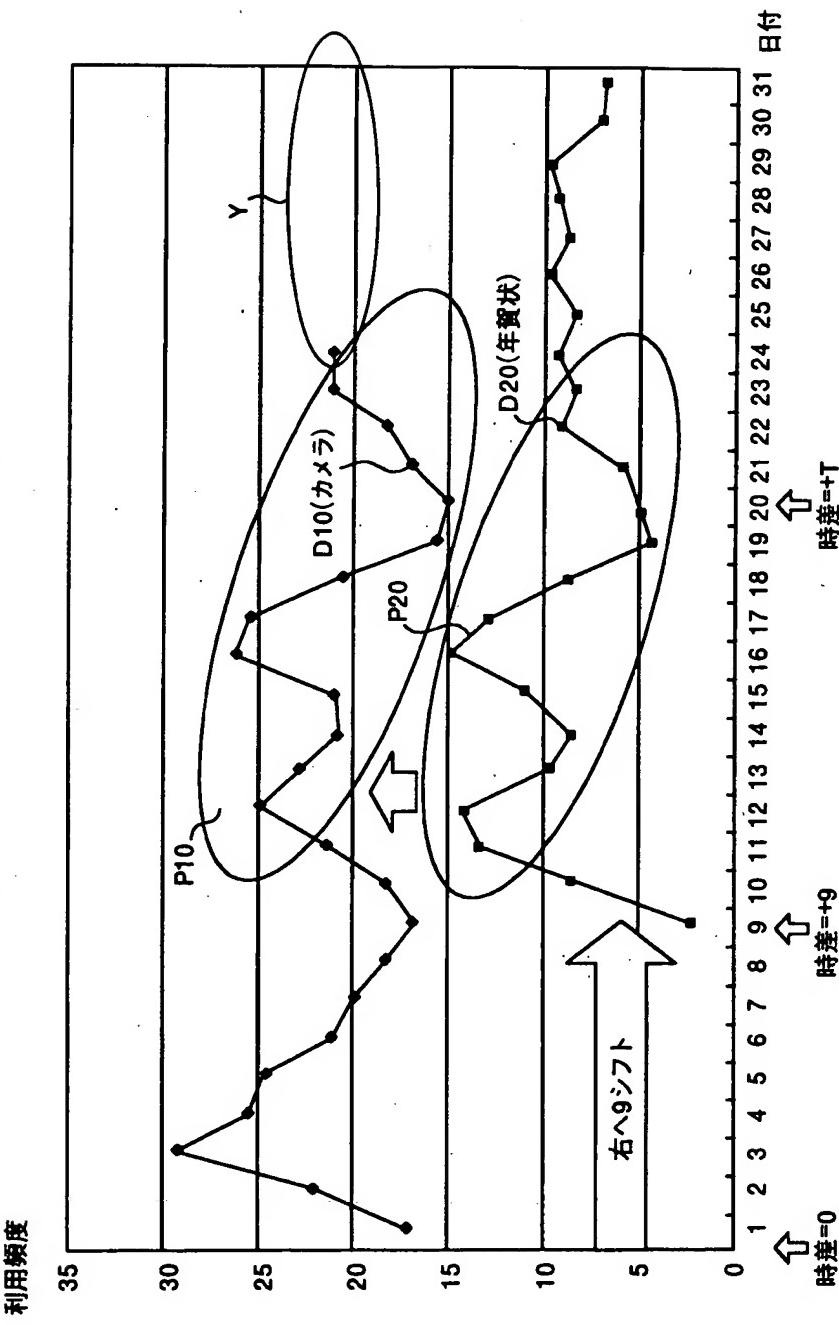
【図5】

図1に示した比較結果データCの一例を示す図

第1の時系列データ カメラ	第2の時系列データ 年賀状	相関係数 0.1	時差 -T
カメラ	年賀状	0.16	0
カメラ	年賀状	0.68	+9
カメラ	年賀状	0.2	+T
カメラ	ラジオ	0.05	-T
...

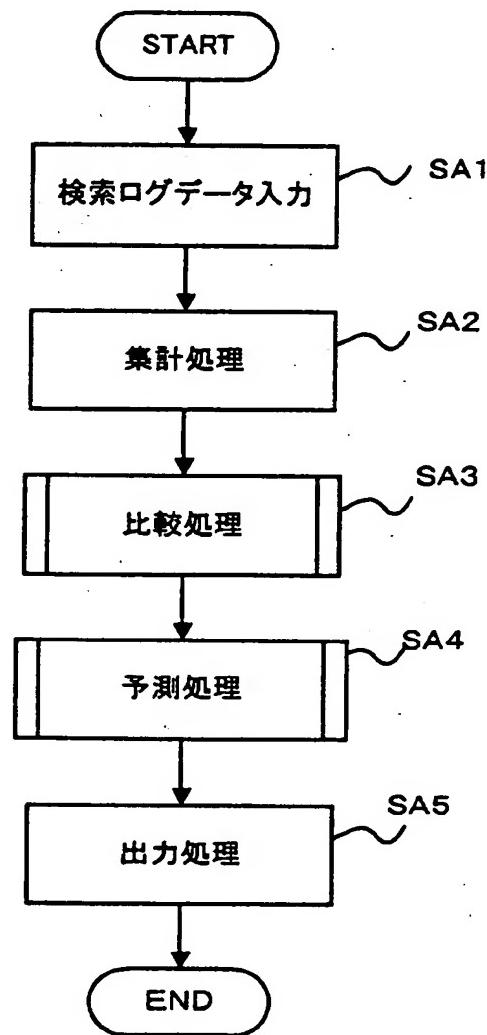
【図6】

実施の形態1における右シフトを説明する図



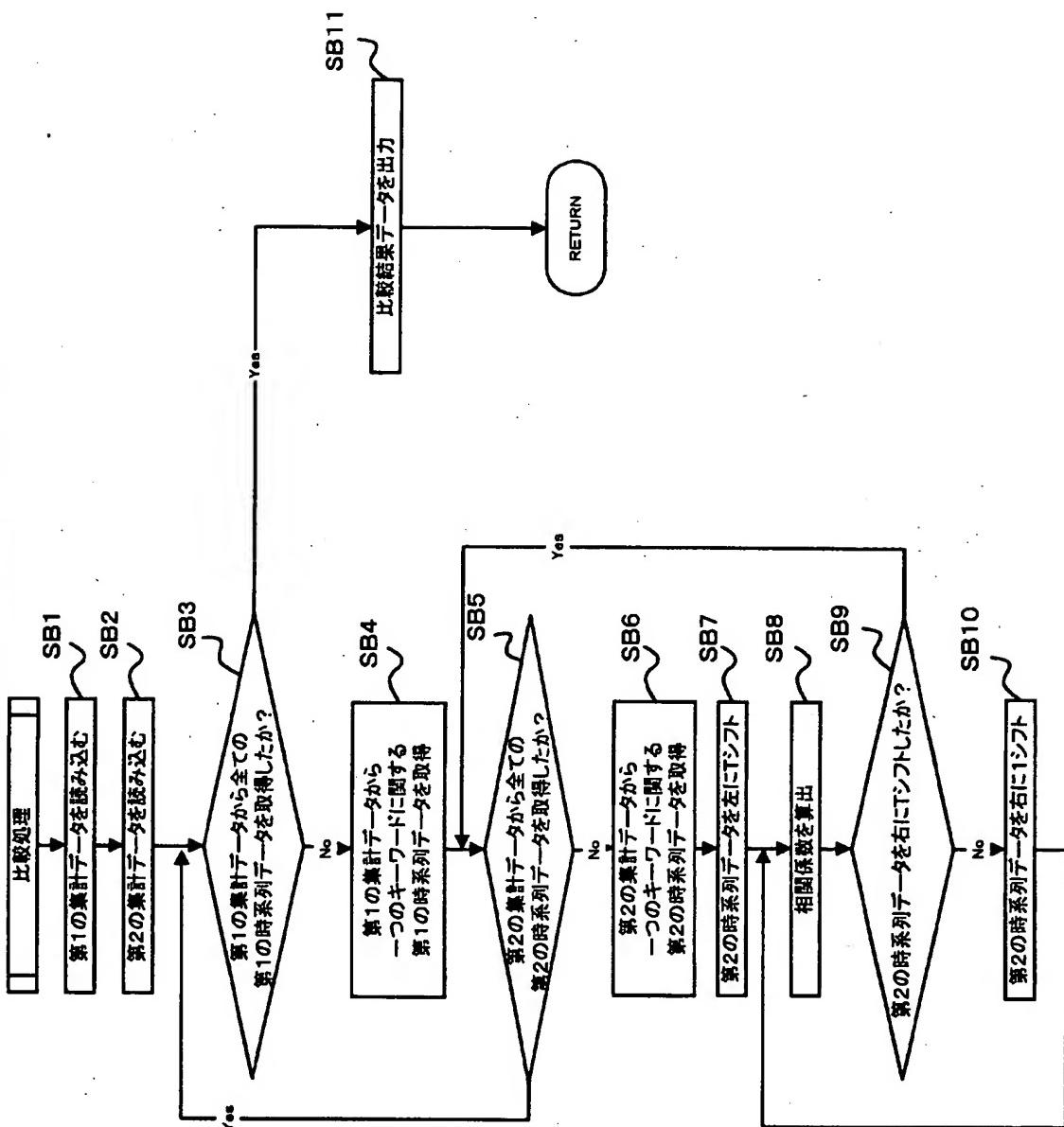
【図7】

実施の形態1の動作を説明するフローチャート



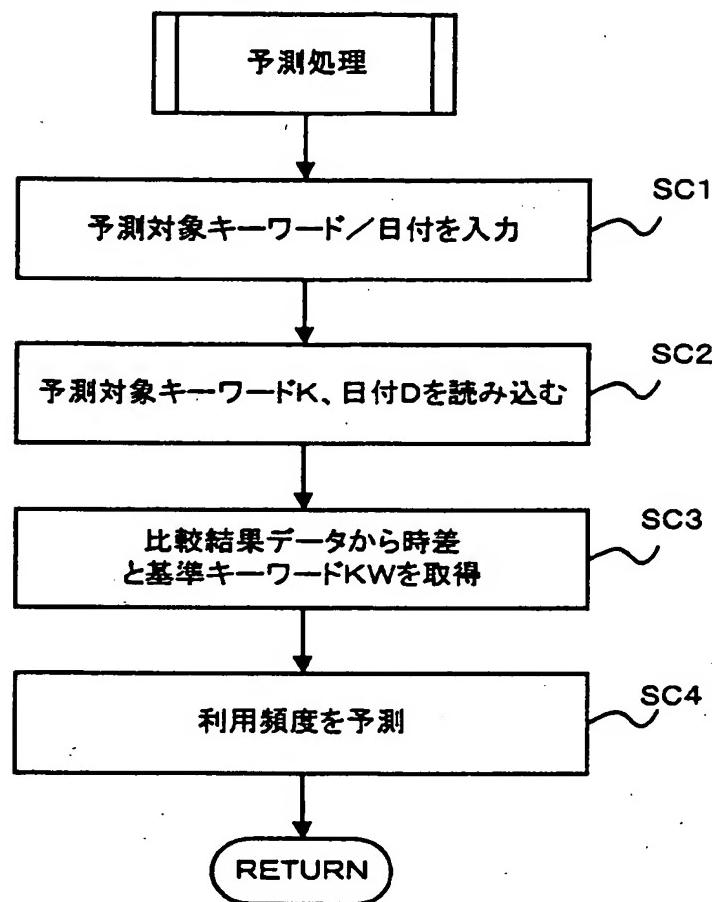
【図8】

図7に示した比較処理を説明するフローチャート



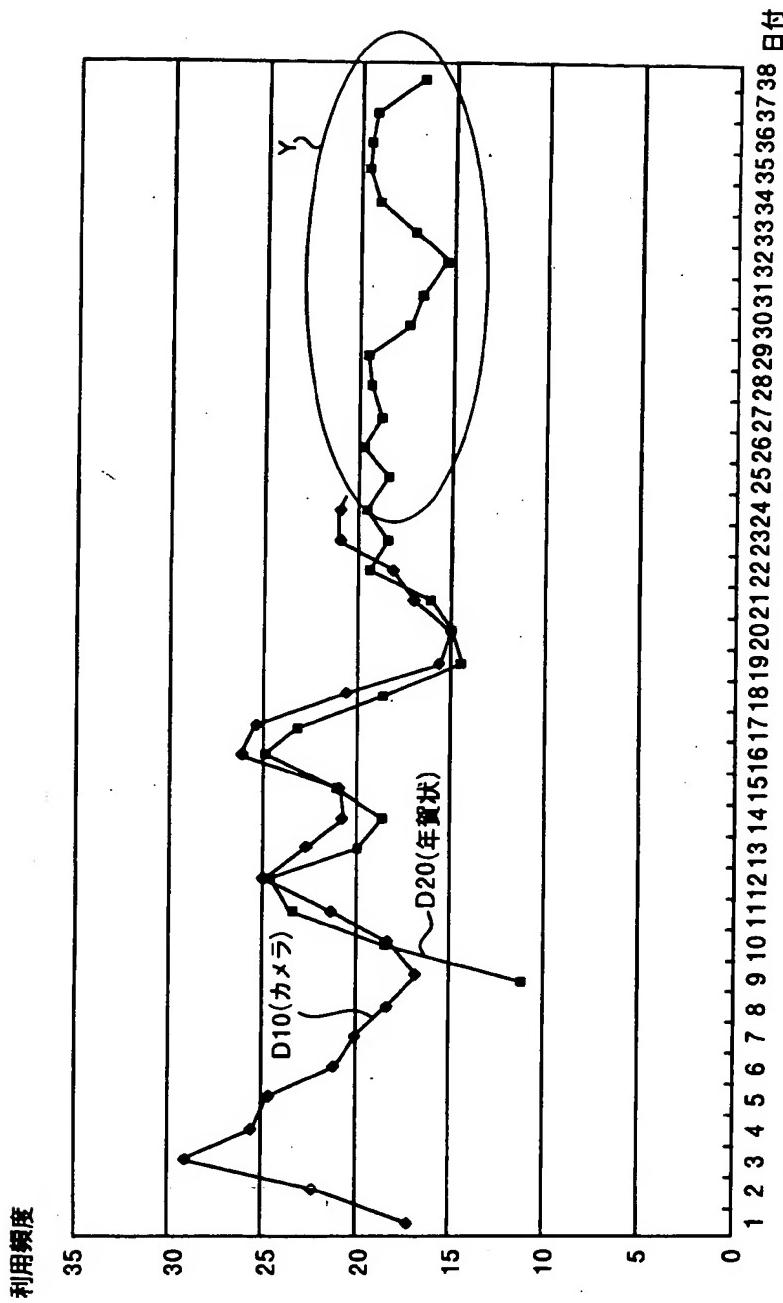
【図9】

図7に示した予測処理を説明するフローチャート



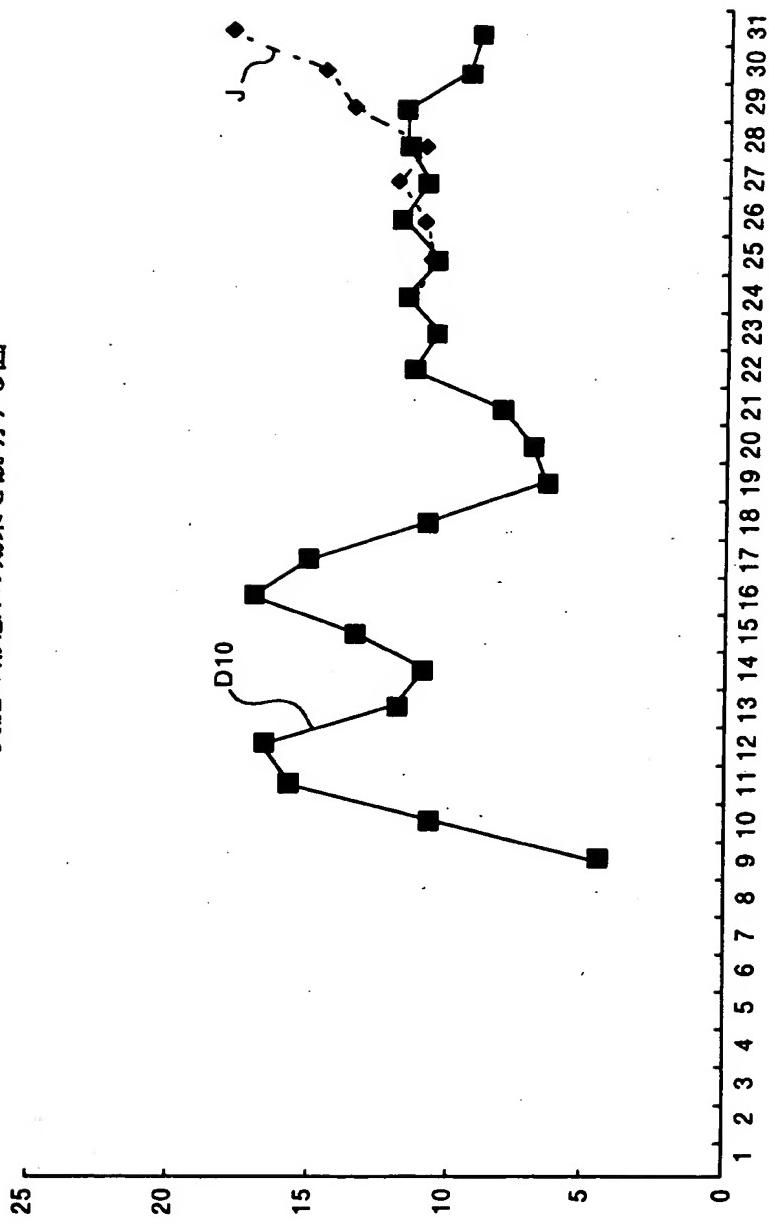
【図10】

実施の形態1の動作を説明する図

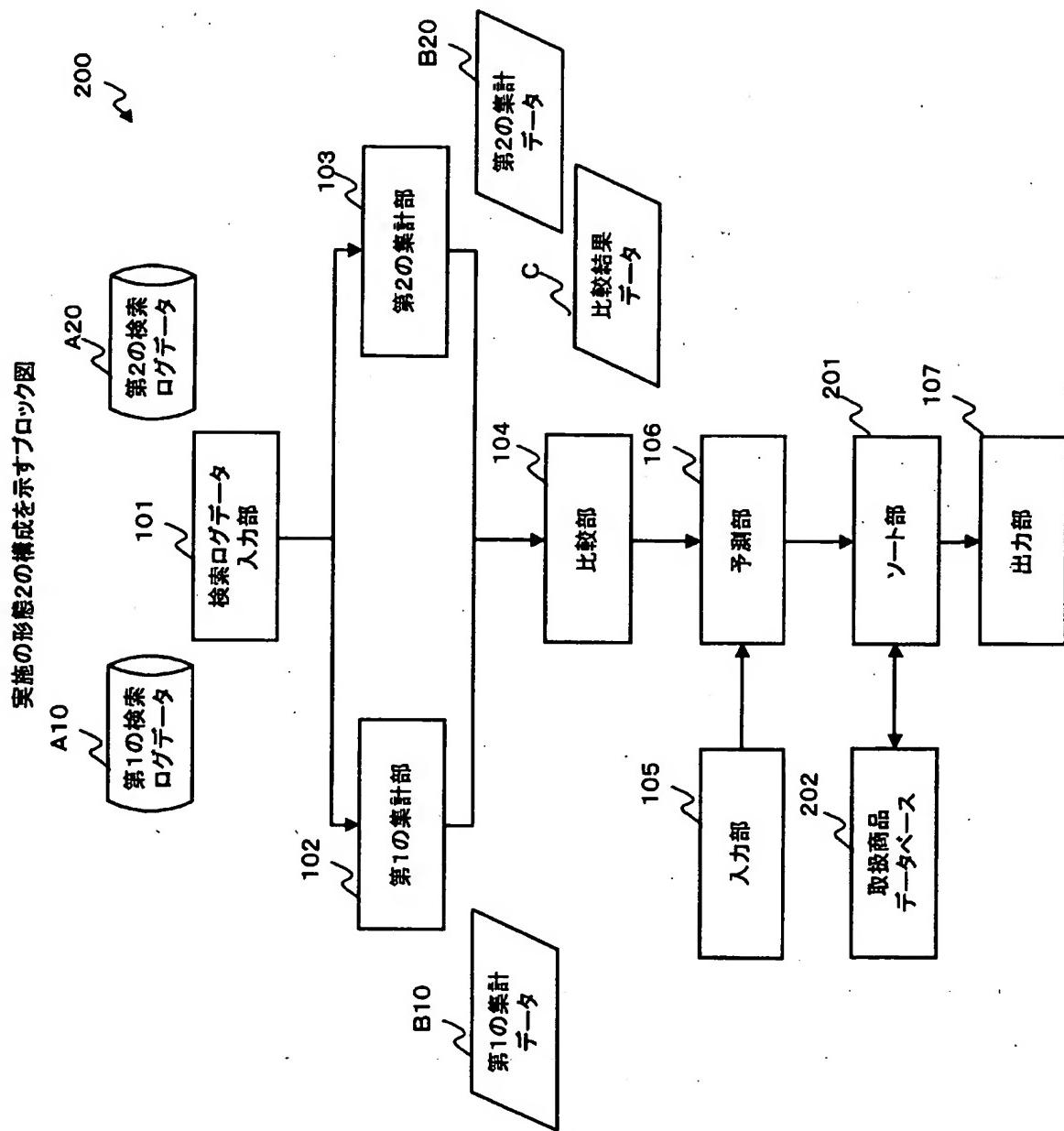


【図11】

実施の形態1の効果を説明する図

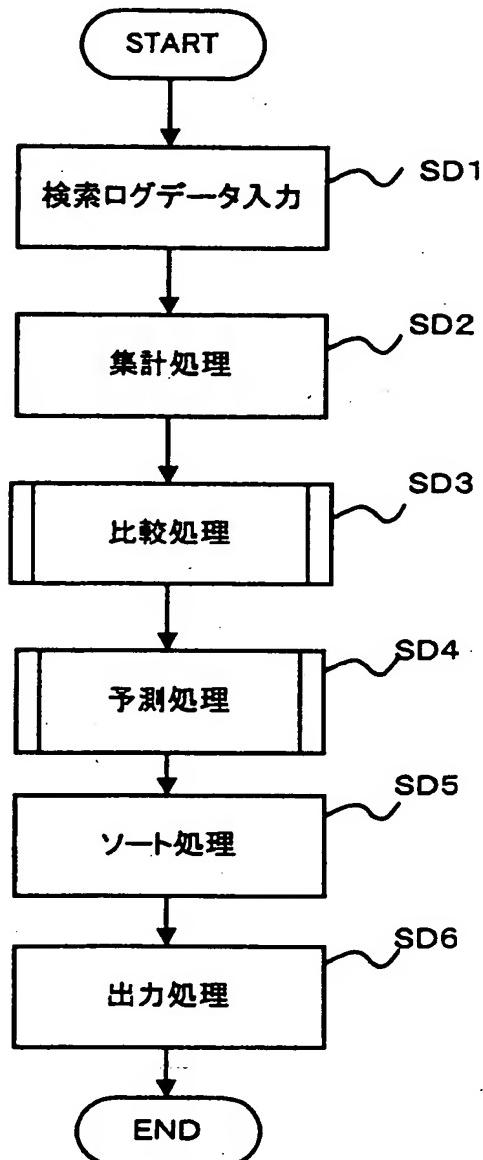


【図12】



【図13】

実施の形態2の動作を説明するフローチャート



【図14】

図13に示したシート処理を説明する図

E1

取扱商品	予測利用頻度
デジタルカメラ	25
カメラ	36
フィルム	12
三脚	3
フレンシュ	7
コンパクトカメラ	20
一眼レフ	2

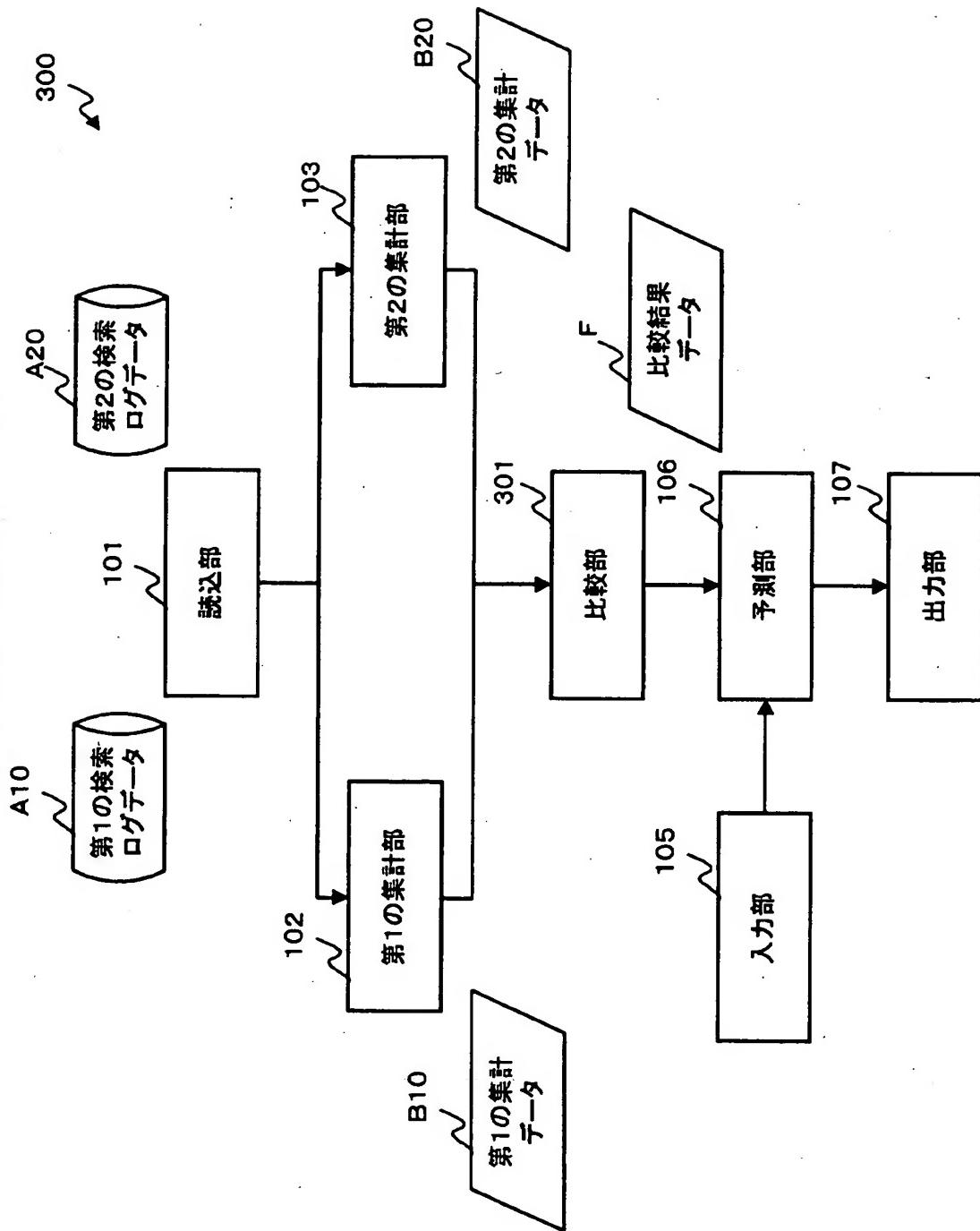
ソート

E2

取扱商品	予測利用頻度
カメラ	36
デジタルカメラ	25
コンパクトカメラ	20
フィルム	12
フレンシュ	7
三脚	3
一眼レフ	2

【図15】

実施の形態3の構成を示すプロック図



【図16】

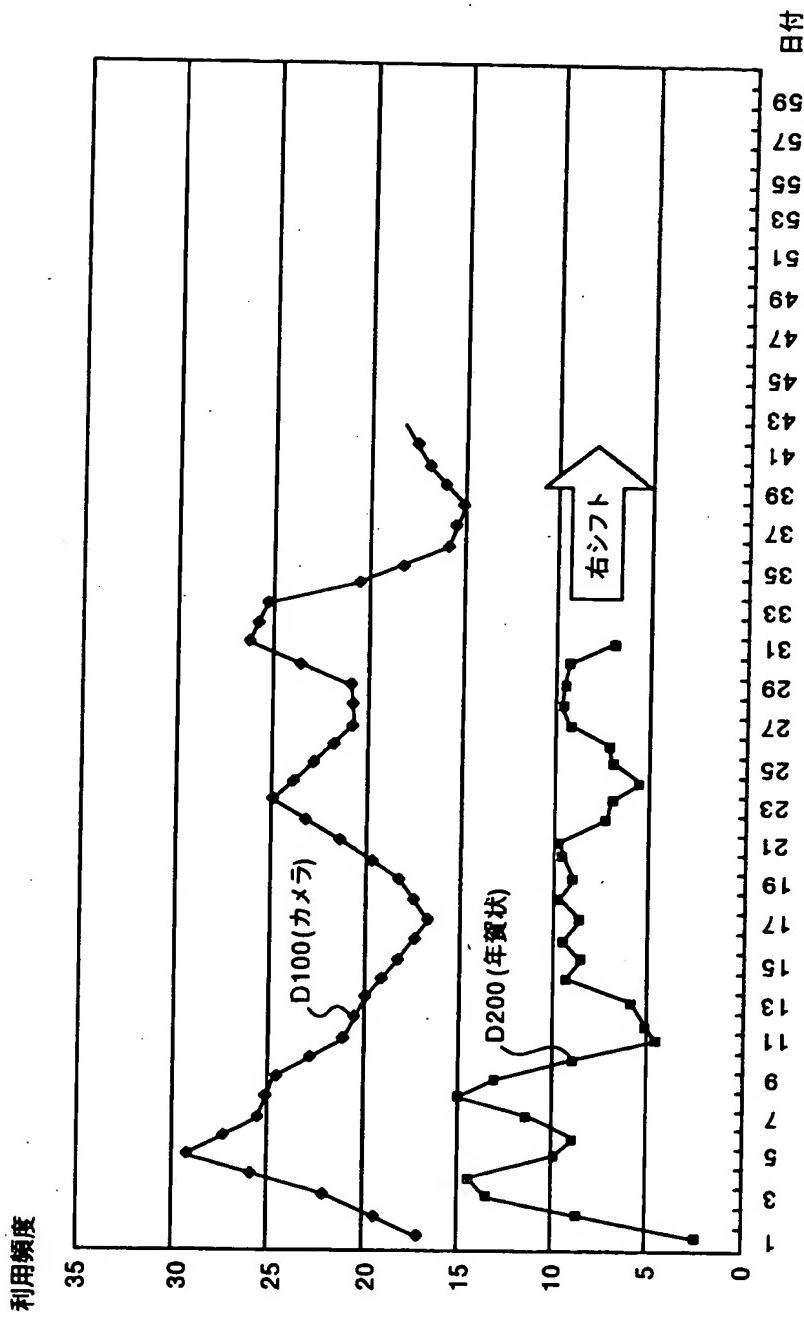
図15に示した比較結果データFの一例を示す図

F

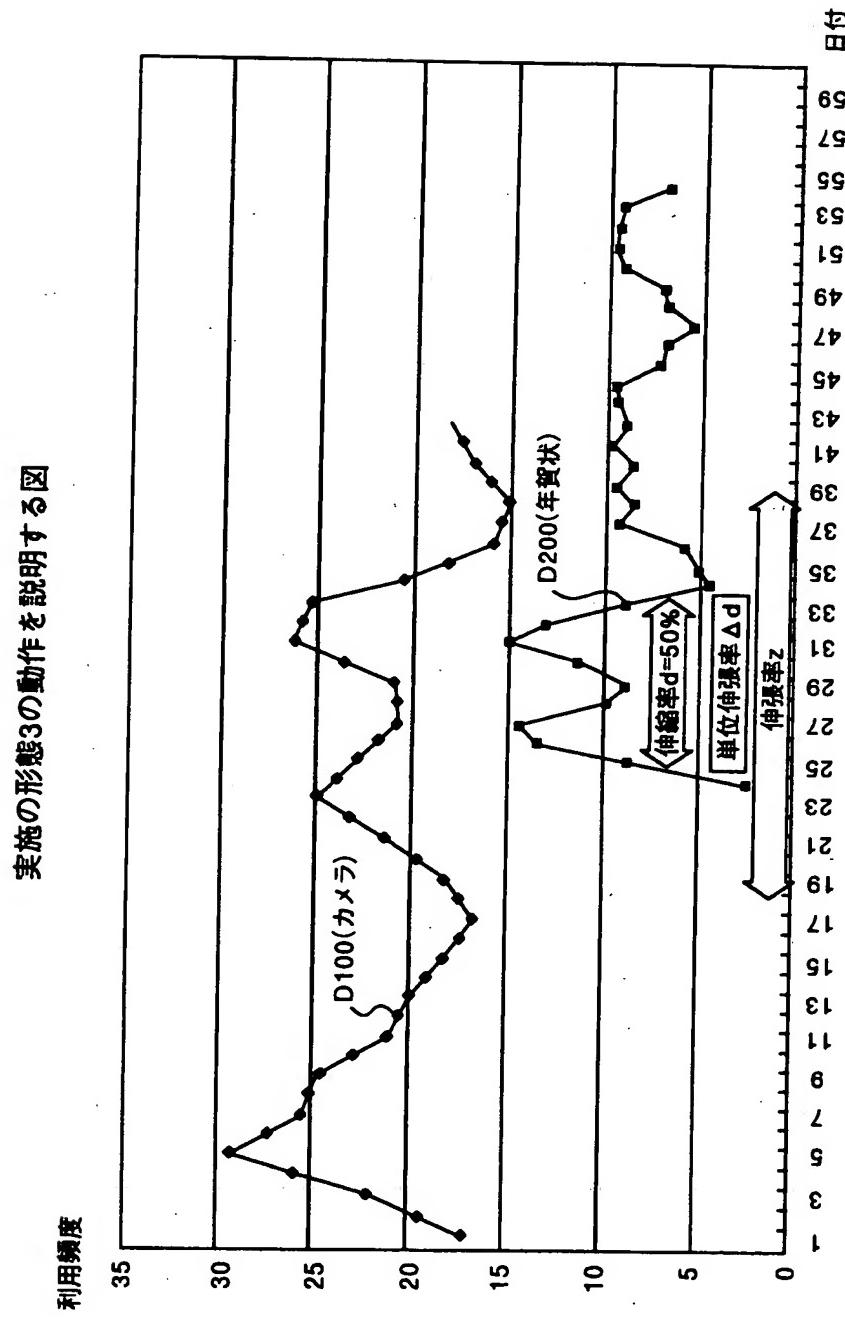
第1の時系列データ	第2の時系列データ	相関係数	時差	伸縮率
カメラ	年賀状	0.1	-T	50%
カメラ	年賀状	0.16	-T	100%
カメラ	年賀状	0.68	-T	200%
カメラ	年賀状	0.2	-T+1	50%

【図17】

実施の形態3の動作を説明する図

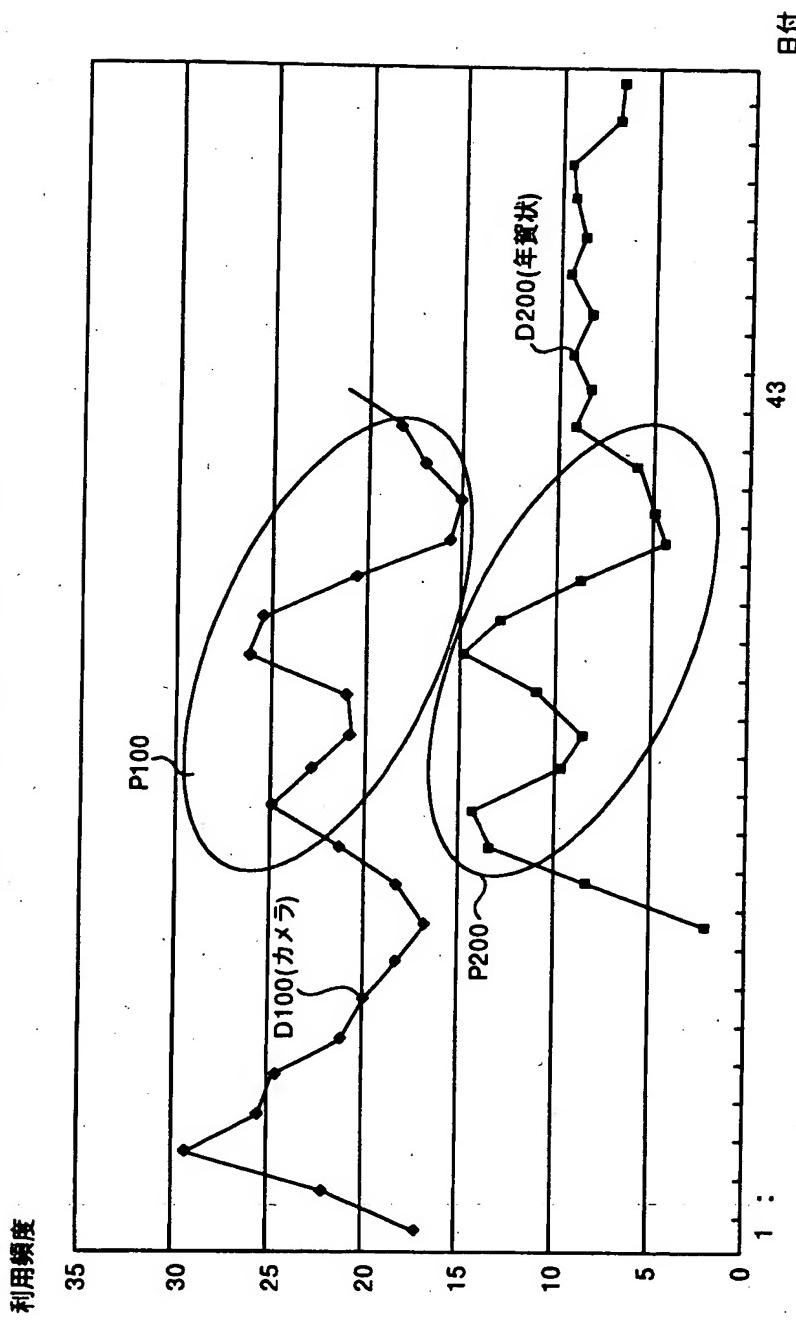


【図18】



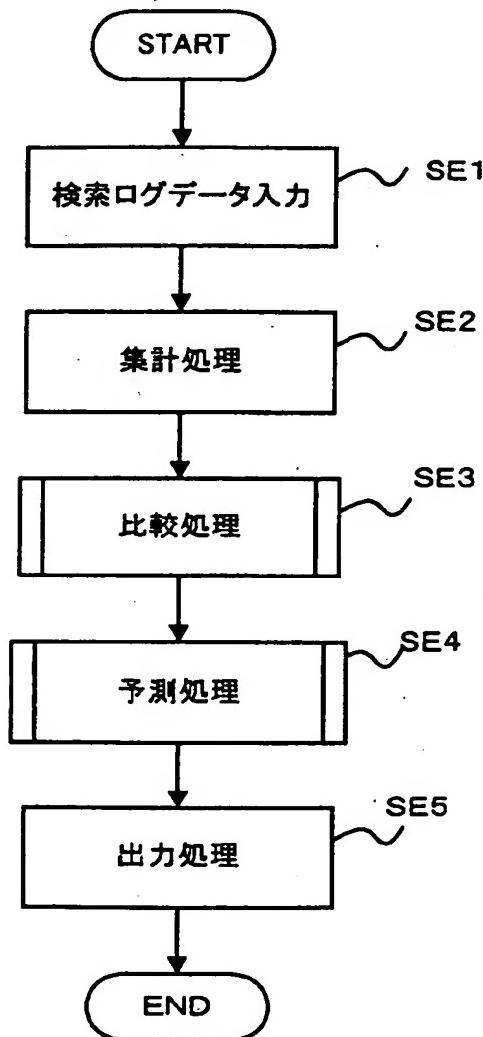
【図19】

実施の形態3の動作を説明する図



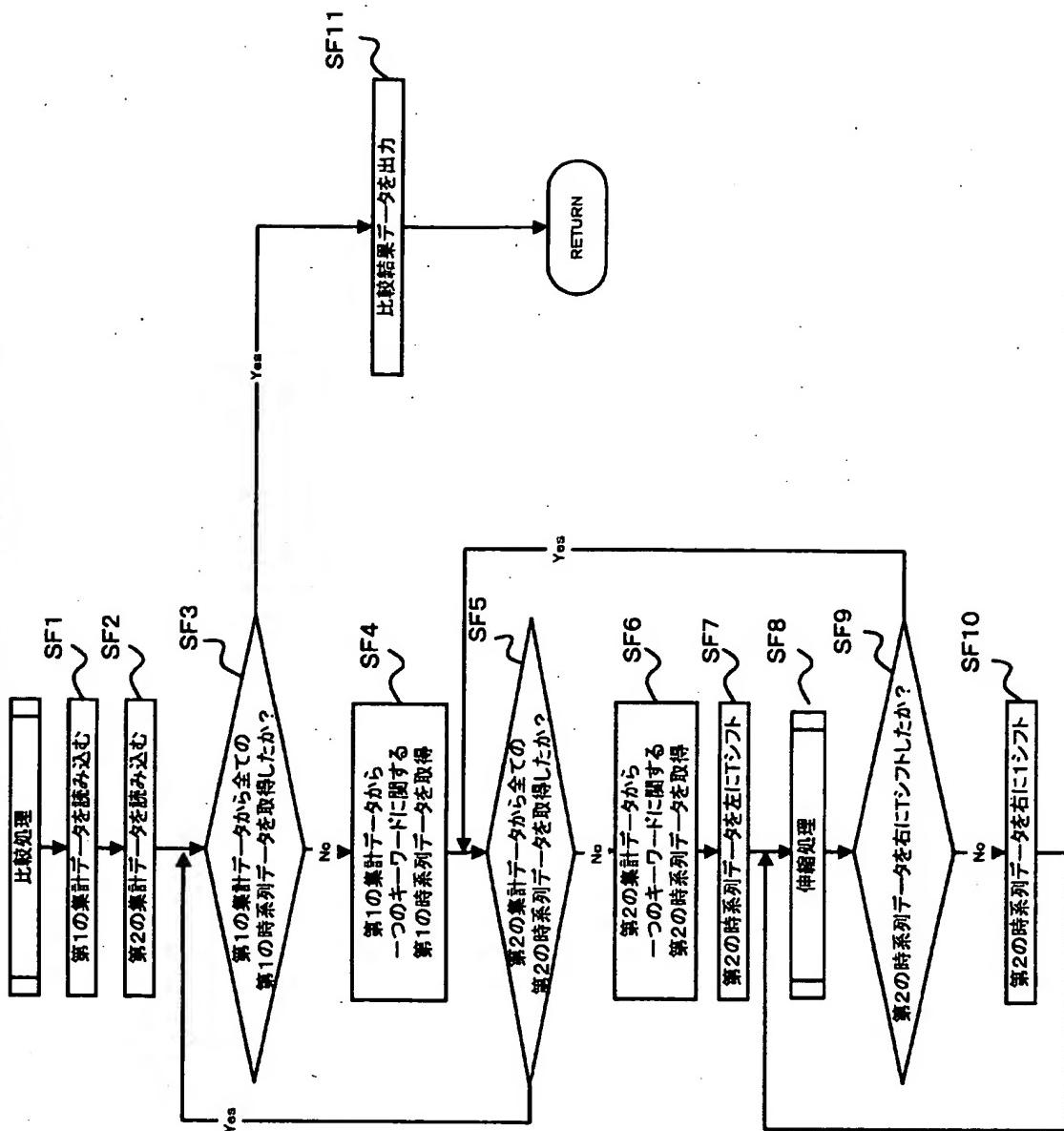
【図20】

実施の形態3の動作を説明するフローチャート



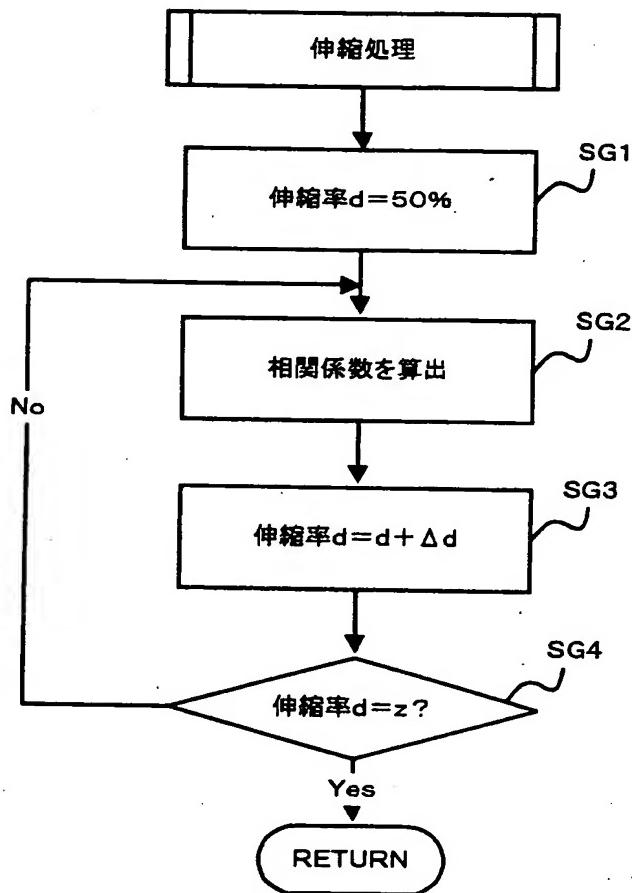
【図21】

図20に示した比較処理を説明するフローチャート



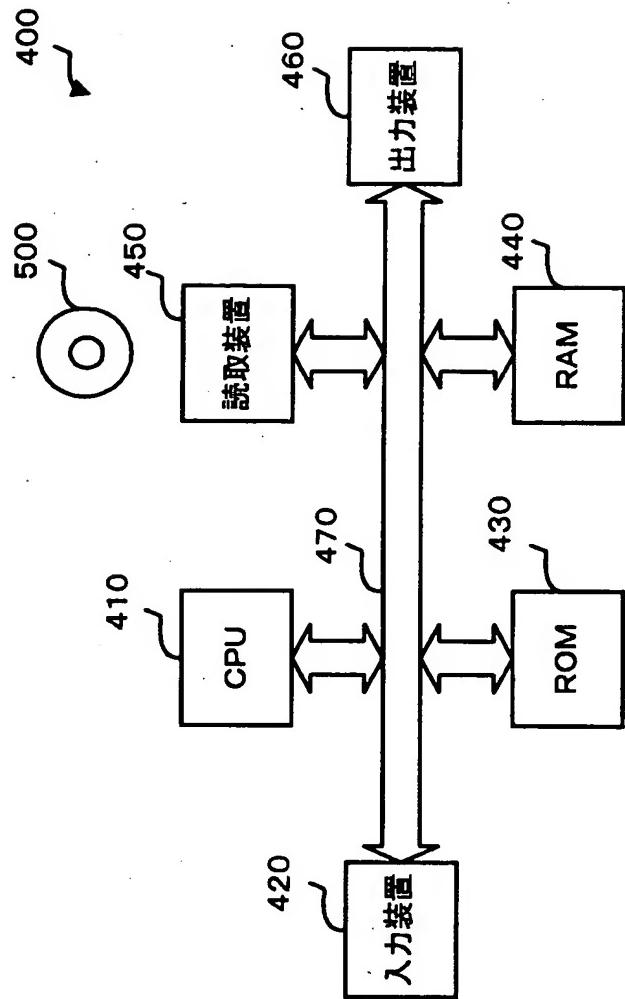
【図22】

図21に示した伸縮処理を説明するフローチャート



【図23】

実施の形態1～3の変形例の構成を示すブロック図



【書類名】 要約書

【要約】

【課題】 情報の利用頻度を高精度で予測すること。

【解決手段】 インターネットのキーワード検索で入力されるキーワードとしての「カメラ」の利用頻度に関する第1の時系列データD10を基準として、別のキーワードとしての「年賀状」の利用頻度に関する第2の時系列データD20を、時間軸上で、順次、所定時間単位でシフトさせ、所定時間単位毎に、第1の時系列データD10と第2の時系列データD20との相関係数を算出する比較部と、比較部により算出された複数の相関係数のうち、最も高い値の相関係数に対応する第1の時系列データD10と第2の時系列データD20との組を特定し、第2の時系列データD20に基づいて、組を構成する第1の時系列データD10の利用頻度を予測する予測部とを備えている。

【選択図】 図4

特2001-153116

出願人履歴情報

識別番号 [000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号
氏 名 富士通株式会社